

Beatriz López
Pau Herrero
Clare Martin (Eds.)



Artificial Intelligence for Diabetes

1st ECAI Workshop on Artificial intelligence for Diabetes
at the
22nd European Conference on Artificial Intelligence (ECAI 2016)

30 August 2016, The Hague, Holland

Proceedings

Volume Editors

Beatriz López
Department of electric, Electronic and Automatic Engineering
University of Girona

Pau Herrero
Department of Electrical and Electronic Engineering
Imperial College London

Clare Martin
Department of Computing and Communication Technologies
Oxford Brookes University

This volume contains the proceedings of the workshop “Artificial Intelligence for Diabetes” at ECAI 20016

Organising Committee

Beatriz López, *University of Girona, Spain. Email: beatriz.lopez@udg.edu*

Pau Herrero, *Imperial College, London, UK. Email: p.herrero-vinias@imperial.ac.uk*

Clare Martin, *Oxford Brookes University, UK. Email: cemartin@brookes.ac.uk*

Program Committee:

Arantza Aldea, *Oxford Brookes University, United Kingdom*

Stefano Bromuri, *Open University, Netherlands*

Lucian Nita, *Romsoft SRL, Romania*

Michael Ignaz Schumacher, *Institute of Business Information Systems, University of Applied Sciences, Western Switzerland*

Cindy Marling, *Ohio University, USA*

Eva Armengol, *Artificial Intelligence Research Institute (IIIA), Spain*

David Riaño, *Universitat Rovira i Virgili, Spain*

Nick Oliver, *Imperial College London, United Kingdom*

Pantelis Giorgiou, *Imperial College London, United Kingdom*

Jorge Bondia, *Instituto Universitario de Automática e Informática Industrial, Departamentp de Ingeniería de Sistemas y Automática, Universidad Politécnica de Valencia, Spain*

Ozgür Kafali, *NC State University, USA*

Magí Lluch-Ariet, *Eurecat Technology Centre, Catalonia.*

Sponsors



The PEPPER project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 689810.

Editor's preface

The complexity of diabetes prognosis and management has lead Artificial Intelligence (AI) to become a key technology to provide solutions that empower both patients and caregivers in their everyday life. Several publicly-funded projects have been carried out, such as: EMPOWER, MOBIGUIDE, COMMODITY12 EU, DIADVISOR, DIABEO, and the recently launched PEPPER project. However, there is still a lot of work left to be done. The aim of this workshop is to assimilate lessons learned, and discuss future work, as a first step towards finding definitive, compatible and complementary AI tools for people dealing with diabetes.

The AID workshop will therefore facilitate discussion among different researchers actively engaged in finding AI-based solutions to problems associated with diabetes. Ten papers have been accepted, which represent a sample of the latest research in the area by several research groups. The final session of the workshop schedule is designated for discussion of the next steps to keep the community engaged and growing, including the proposal of new collaborative projects. We hope that you will enjoy the workshop and join the community in the forthcoming events that stem from it.

The Organizing Committee
Beatriz López (University of Girona, Spain)
Pau Herrero (Imperial College of London, UK)
Clare Martin (Oxford Brookes University, UK)

The Hague, Netherlands
August 30th, 2016

Keynote

Speaker: Prof. Riccardo Bellazzi, *Dipartimento di Ingegneria Industriale e dell'Informazione, Università degli Studi di Pavia, Italy*

Title: *"Artificial Intelligence in Diabetes Mellitus management: advanced strategies for a complex disease"*

Abstract: Diabetes Mellitus, due to its multi-faceted, dynamics and data-intensive nature, is a paradigmatic disease for the application of AI-based approaches, including rule-based, case-based and model-based reasoning, machine learning and visual analytics. Starting from the lessons learned from past and current research projects, the talk will discuss some future research directions for the integration of AI into the clinical management of Diabetes.

Table of Contents

PEPPER: Patient Empowerment Through Predictive Personalised Decision Support	8
Herrero, Pau and López, Beatriz and Martin, Clare	
Enhancing an Artificial Pancreas with an Adaptive Bolus Calculator based on Case-Based Reasoning.....	10
Herrero, Pau and Bondia, Jorge and Pesl, Peter and Oliver, Nick and Georgiou, Pantelis	
Temporal Case-Based Reasoning for Bolus Decision Support	14
Brown, Daniel and Harrison, Rachel and Martin, Clare and Bayley, Ian	
Monitoring Patients with Diabetes Using Wearable Sensors: Predicting Glycaemias Using ECG and Respiration Rate	18
Cvetković, Božidara and Pangerc, Urška and Gradišek, Anton and Luštrek, Mitja	
Developing a Motivational System to Manage Physical Activity for Type 2 Diabetes	22
Alfaifi, Yousef and Grasso, Floriana and Tamma, Valentina	
Increasing Transparency of Recommender Systems for Type 1 Diabetes Patients	26
Vargheese, John Paul and Harrison, Rachel and Munoz Balbontin, Mireya and Aldea, Arantza and Brown, Daniel	
Assessment of Diabetic Complications Based on Series of rRecords	28
Armengol, Eva	
D1NAMO, A Personal Health System for Glycemic Events Detection	31
Dubosson, Fabien and Bromuri, Stefano and Ranvier, Jean-and Eudes, Schumacher, Michael	
Ontologies for Social, Cognitive and Affective Agent-Based Support of Child's Diabetes Self-Management	35
Neerincx, Mark A. and Kaptein, Frank and Bekkum, Michael A. van and Krieger, Hans-Ulrich and Kiefer, Bernd and Peters, Rifca and Broekens, Joost and Demiris, Yiannis and Sapelli, Maya	
Handling Missing Phenotype Data with Random Forests for Diabetes Risk Prognosis	39
López, Beatriz and Viñas, Ramon and Torrent-Fontbona, Ferran and Fernandez-Real, José Manuel	

PEPPER: Patient Empowerment Through Predictive Personalised Decision Support

Pau Herrero,¹ Beatriz López² and Clare Martin³

Abstract. PEPPER is a newly-launched three-year research project, funded by the EU Horizon 2020 Framework. It will create a portable personalised decision support system to empower individuals on insulin therapy to self-manage their condition. PEPPER employs Case-Based Reasoning to advise about insulin bolus doses, drawing on various sources of physiological, lifestyle, environmental and social data. It also uses a Model-Based Reasoning approach to maximise users' safety. The system will be integrated with an unobtrusive insulin patch pump and has a patient-centric development approach in order to improve patient self-efficacy and adherence to treatment.

1 INTRODUCTION

Type 1 diabetes (T1D) is a chronic disease caused by an autoimmune destruction of the pancreatic beta cells. This leaves the body unable to produce the insulin needed to regulate blood glucose levels. The condition is usually controlled through multiple daily injections (MDI) of insulin to mimic the natural insulin secretion of a healthy pancreas. Alternatively, some people are treated with continuous subcutaneous insulin infusion (CSII) via a wearable pump. In both cases the size of each insulin doses are chosen by the individual.

Decision support tools exist to support this process, such as insulin bolus calculators that use simple mathematical formulae based on metabolic parameters (i.e. insulin-to-carbohydrate ratio and insulin sensitivity factor) and an estimation of the active insulin from previous doses. Such tools are integrated into most insulin pumps [9], and some glucose meters. There is also an increasing adoption of decision support tools implemented on mobile devices [10], often in conjunction with remote data storage in the cloud, though few are approved by regulatory bodies such as the FDA. Some gather inputs via wearable sensors (i.e. continuous glucose monitors), but most of them rely on manual input. In practice, the latter are rarely used because most people with T1D find the process tedious and refuse to interact with such systems [10, 2]. Hence a guiding design principle for PEPPER is that wherever possible data is collected automatically, via wearable technology. The information collected by the sensors is managed by a Case-Based Reasoning (CBR) module to provide personalised insulin recommendations, while a second Model-Based Reasoning (MBR) module is used to maximise users' safety.

2 SYSTEM OVERVIEW

The PEPPER system shown in Figure 1 offers insulin dosing advice that is highly adaptive to the insulin needs of individuals by using

a CBR approach. It also guarantees individuals' safety by means of a MBR approach that includes predictive glucose alarms, automatic insulin suspension, carbohydrate recommendations and fault diagnosis. PEPPER offers a dual architecture to cater for both MDI or CSII treatment, the latter via the unobtrusive Cellnovo patch-pump (Cellnovo Ltd., UK). In both cases, the patient periodically wears a continuous glucose monitor (CGM) used to automatically evaluate glucose outcomes. An activity monitor, such the one integrated in the Cellnovo pump or a commercially available one (e.g. Fitbit), is included to determine physical activity automatically. Data from a capillary blood glucose meter is periodically gathered to calibrate the CGM or to be used in case CGM data is not available. Additional data such as food intake, alcohol consumption, hormonal cycles are input through the user interface of the handheld unit (smartphone or Cellnovo handset). All inputs are then fed to the CBR engine on the handheld unit, and used to calculate the corresponding insulin dose. The dose is then displayed for the user to accept or decline. If the recommendation is accepted, the unit wirelessly sends the corresponding command to the insulin pump, or the user manually injects the bolus using an insulin pen. In addition, the safety module triggers alarms to alert the user about predicted hypo- and hyperglycaemic events. In the case of impending hypoglycaemia, the system also recommends a personalised amount of carbohydrates to consume to eliminate hypoglycaemia and avoid rebound hyperglycaemia. It also suspends insulin delivery for pump users when glucose levels are forecast to be too low. If potentially dangerous events are not properly addressed by the subject, automatic alarms can be sent via an SMS service to the expert team and selected carers. When network connectivity is available, the handheld unit sends the recorded data to a remote secure server. Data is presented in meaningful visualisations and analysed periodically to find non-optimal glucose patterns.

2.1 Case-Based Reasoning for Insulin Dosing

Case-Based Reasoning (CBR) is a consolidated artificial intelligence technique, extensively applied in medicine, that tries to solve newly encountered problems by applying solutions learned from similar problems encountered in the past. In CBR, past situations are stored in *cases*, which represent knowledge related to the various aspects of the situation. The CBR cycle consists of four steps: *Retrieve* the most similar case or cases; *Reuse* the information in that case to solve the problem; *Revise* the proposed solution; *Retain* the parts of this experience likely to be useful for future problem solving [1].

The first project to use CBR to recommend changes in insulin therapy for T1D management was the T-IDDM project [3], where it was integrated with rule-based reasoning and a probabilistic model of the effects of insulin on blood glucose levels. More recently, the IDSDM

¹ Imperial College London, email: pherrero@imperial.ac.uk

² University of Girona, email: beatriz.lopez@udg.edu

³ Oxford Brookes University, email: cemartin@brookes.ac.uk

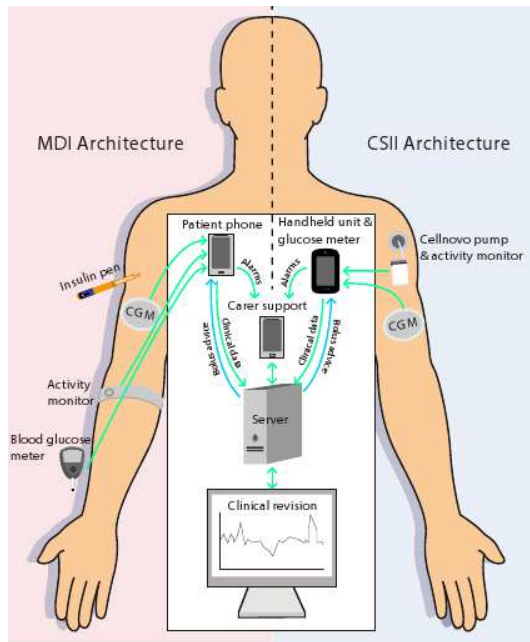


Figure 1. PEPPER architecture.

project [11] used CBR as the primary reasoning modality in a decision support tool for patients on insulin pump therapy, and introduced other factors into the calculations, such as life events that can influence blood glucose levels. However, both projects were intended for use by clinicians as opposed to the individuals with diabetes.

In PEPPER, the CBR cycle is divided into two parts: the local and remote. The local part runs on the handheld unit and the remote part on a server. Both parts contain a case-base and periodically the local case-base is synchronised with the remote case-base. The evaluation step of the CBR cycle occurs on the server and requires approval by an expert clinician before a new case is incorporated to the case-base. The CBR parameters include CGM and capillary glucose data, physical activity, time, location, basal insulin, hormone cycle, stress, alcohol, meal composition, and sleep. Most of these parameters are automatically collected (or calculated) by the handset unit. Exceptions include alcohol consumption, meal composition and hormone cycles, which need to be manually inputted. A prototype version of the algorithm has already been implemented and successfully tested *in silico* [8] and in subsequent pilot studies [12]. PEPPER builds on this prototype and further improves it by including more parameters and automatising their recording.

2.2 Model-Based Reasoning for Safety

Model-Based Reasoning (MBR) is defined as the interaction of observation and prediction [5]. On the one hand, there is the actual system (e.g. T1D subject) whose behaviour can be observed; on the other hand, there is the model of the system from which predictions (e.g. glucose levels) can be made. Assuming that the models are correct, any discrepancy found between observations and predictions are defaults on the device (e.g. CGM or pump fault). MBR techniques have been previously proposed in the context of diabetes technology to constrain insulin delivery by an artificial pancreas [4], predict hypoglycaemic events [6] and detect CGM and insulin pump faults [7]. PEPPER leverages these techniques to build a system that guarantees safety of the user at any time. In addition, it incorporates an adaptive carbohydrate recommender system to prevent hypoglycaemic events.

3 CONCLUSION

The PEPPER system provides a portable personalised decision support system for insulin dosing that combines data from multiple sources such as body-worn sensors and manual inputs. The Case-Based Reasoning module is designed to provide a personalised insulin dose which adapts over time. A Model-Based Reasoning module is designed to maximise safety through prediction of adverse events and the detection of faults. PEPPER is being developed using a patient-centric approach in order to improve patient self-efficacy and adherence to treatment. The software development will adhere to international standards including those that apply to security and interoperability. The final system will be tested *in silico* before being clinically validated over a 6-month non-randomised open-label ambulatory trial.

ACKNOWLEDGEMENTS

Thanks to the PEPPER team: A. Aldea, D. Brown, D. Duce, JM Fernández-Real, P. Georgiou, J. González López, R. Harrison, B. Innocenti, J. Masoud, L. Nita, N. Oliver, P. Pesl, R. Petite, M. Reddy, J. Shapley, F. Torrent, C. Toumazou, JP Vargheese, and M. Waite. This project has received funding from the EU Horizon 2020 research and innovation programme under grant agreement No 689810.

REFERENCES

- [1] Agnar Aamodt and Enric Plaza, 'Case-based reasoning: Foundational issues, methodological variations, and system approaches', *AI communications*, 7(1), 39–59, (1994).
- [2] Eirik Årsand, Dag Helge Frøisland, Stein Olav Skrøvseth, Taridzo Chomutare, Naoe Tatara, Gunnar Hartvigsen, and James T Tufano, 'Mobile health applications to assist patients with diabetes: lessons learned and design implications', *Journal of diabetes science and technology*, 6(5), 1197–1206, (2012).
- [3] Riccardo Bellazzi, Cristiana Larizza, Stefania Montani, Alberto Riva, Mario Stefanelli, Giuseppe d'Annunzio, R Lorini, Enrique J Gómez, E Hernando, Eulalia Brugués, et al., 'A telemedicine support for diabetes management: the t-iddm project', *Computer methods and programs in biomedicine*, 69(2), 147–161, (2002).
- [4] Ellingsen C., Dassau E., Zisser H., and et al., 'Safety constraints in an artificial pancreatic cell: An implementation of model predictive control with insulin on board', *Journal of diabetes science and technology*, 3(3), 536–544, (2009).
- [5] Randall Davis and Walter C. Hamscher, 'Model-based reasoning: Troubleshooting', in *Exploring Artificial Intelligence*, pp. 297–346. Morgan Kaufmann Publishers Inc., (1988).
- [6] RA Harvey, E Dassau, H Zisser, DE Seborg, L Jovanovic, and Doyle FJ., 'Design of the health monitoring system for the artificial pancreas: Low glucose prediction module.', *Journal of Diabetes Science and Technology*, 6(6), 1345–1354, (2012).
- [7] P Herrero, R Calm, J Veh, J Armengol, P Georgiou, N Oliver, and Tomazou C., 'Robust fault detection system for insulin pump therapy using continuous glucose monitoring.', *Journal of Diabetes Science and Technology*, 6(5), 1131–41, (2012).
- [8] Pau Herrero, Peter Pesl, Monika Reddy, Nick Oliver, Pantelis Georgiou, and Christofer Toumazou, 'Advanced insulin bolus advisor based on run-to-run control and case-based reasoning', *Biomedical and Health Informatics, IEEE Journal of*, 19(3), 1087–1096, (2015).
- [9] David C Klonoff, 'The current status of bolus calculator decision-support software', *J Diabetes Sci Technol*, 6(5), 990–994, (2012).
- [10] David C Klonoff, 'Telemedicine for diabetes current and future trends', *Journal of diabetes science and technology*, 10(1), 3–5, (2016).
- [11] Cindy Marling, Jay Shubrook, and Frank Schwartz, 'Case-based decision support for patients with type 1 diabetes on insulin pump therapy', in *Advances in Case-Based Reasoning*, 325–339, Springer, (2008).
- [12] Monika Reddy, Peter Pesl, Pau Herrero, and et al., 'Clinical safety and feasibility of the advanced bolus calculator for type 1 diabetes based on case-based reasoning: a 6-month randomised single-arm pilot study.', *Diabetes Technol Ther* 2016, (2016).

Enhancing an Artificial Pancreas with an Adaptive Bolus Calculator based on Case-Based Reasoning

Pau Herrero¹, Jorge Bondia², Peter Pesl³, Nick Oliver⁴ and Pantelis Georgiou⁵

Abstract. Current prototypes of closed-loop systems for glucose control in type 1 diabetes mellitus, also referred to as artificial pancreas systems, require a pre-meal insulin bolus to compensate for delays in subcutaneous insulin absorption in order to avoid initial post-prandial hyperglycemia. Most closed-loop systems compute this pre-meal insulin dose by a standard bolus calculation, as is commonly found in insulin pumps. However, the performance of these calculators is limited due to a lack of adaptiveness in front of dynamic changes in insulin requirements. In this paper we present a new technique to automatically adapt the meal-priming bolus within an artificial pancreas based on Case-Based Reasoning and Run-To-Run control. Simulation results showed that using an adaptive meal bolus calculator within a closed-loop control system has the potential to improve glycemic control in type 1 diabetes when compared to its non-adaptive counterpart.

1 Introduction

1.1 Type 1 diabetes mellitus (T1DM)

T1DM is an autoimmune condition characterized by elevated blood glucose levels due to the lack of endogenous insulin production. People with T1DM require exogenous insulin delivery to regulate glucose. Current therapies for T1DM management include the administration of multiple daily injections or continuous insulin infusion with pumps. However, such therapies are still suboptimal and require constant adjustment by the person with T1DM and carers.

1.2 Artificial Pancreas

A closed-loop control system consisting of a continuous glucose sensor, an insulin pump and an algorithm that computes the required insulin dose at any instant, has the potential to improve glucose control in people with T1DM [6]. Ideally, a completely automated closed-loop control system would not require any user intervention, for example to announce meals, and would react in real-time to changes in blood glucose. However, delays in subcutaneous insulin absorption have led many investigators to include the use of a pre-meal insulin bolus within the artificial pancreas (Figure 1). The calculation of such pre-meal insulin bolus is usually done by means of a simple bolus calculator, found in most insulin pumps. However, accurately computing a meal bolus remains a challenging task due to the high

variability of insulin requirements in T1DM and the uncertainty in carbohydrate estimations.

1.3 Adaptive meal-priming bolus

The utilisation of an adaptive meal-priming bolus within an artificial pancreas has previously been proposed by El-Khatib et. al [3] showing some encouraging clinical results relative to an entirely reactive system with no meal-priming boluses. However, this method has the limitation that assumes that carbohydrate intakes are fairly similar every day, which is not always the case. It also does not take into consideration other factors such as exercise, alcohol, stress, weather, hormones, and variation in macronutrient composition. In this paper, we present a novel technique to automatically adjust the meal-priming bolus within an artificial pancreas that overcomes these limitations by allowing the system to consider an estimation of the carbohydrate intake and other parameters affecting glucose outcomes.

2 Methods

The proposed adaptive meal bolus calculator for closed-loop control is based on an existing technique referred to as Advanced Bolus Calculator for Diabetes Management (ABC4D) [2], which has previously been validated tested in clinical trials [10]. ABC4D enhances currently existing bolus calculators by means of a combination of Case-Based Reasoning [1] and Run-To-Run control [8]. Periodic use of continuous glucose monitoring (CGM) data is required in order to perform a retrospective optimization of the bolus calculator parameters. For evaluation purposes, the clinically validated Imperial College Bio-inspired Artificial Pancreas (BiAP) controller was employed [5].

2.1 Insulin Bolus Calculator

A standard insulin bolus calculator is defined by the equation

$$B = \frac{CHO}{ICR} + \frac{(G - G_{sp})}{ISF} - IOB, \quad (1)$$

where B (U) is the total calculated bolus, CHO (g) is the estimated amount of ingested carbohydrates, ICR (g/U) is the insulin-to-carbohydrate-ratio, G (mg/dl) is the measured glucose at meal time, G_{sp} (mg/dl) is the glucose set-point, ISF (mg/dl/U) is the insulin sensitivity factor, and IOB (U) is the insulin-on-board, which represents an estimation of the remaining active insulin in the body. The parameters of a bolus calculator (ICR , ISF) can be manually adjusted based, among other parameters, on the time of the day (i.e. breakfast, lunch, dinner), exercise, stress or variation in hormonal

¹ Imperial College London, email: pherrero@imperial.ac.uk

² Polytechnic University of Valencia, email: jbondia@isa.upv.es

³ Imperial College London, email: peter.pesl@imperial.ac.uk

⁴ Imperial College NHS Trust, email: nick.oliver@imperial.ac.uk

⁵ Imperial College London, email: pantelis@imperial.ac.uk

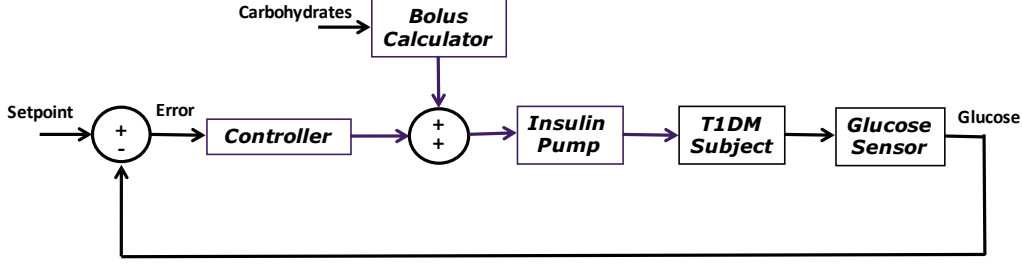


Figure 1. Block diagram of a closed-loop system for glucose control incorporating a meal bolus calculator..

cycles. However, these adjustments are often crude approximations and are rarely revised by the users (subject with T1DM or carer) on a regular basis. In order to provide the required flexibility and adaptability within a bolus calculator to be able to cope with the significant intra-subject variability in T1DM management, a similar approach to the one proposed by Herrero and colleagues [2] was employed. Such approach consist of using Case-Based Reasoning (CBR) to deal with the significant number of case scenarios requiring very different insulin requirements (i.e. solutions) that a person with diabetes has to face. Then, Run-To-Run control is used to automatically revise the parameters of the bolus calculator within the CBR algorithm.

2.2 Case-Based Reasoning (CBR)

CBR is an artificial intelligence problem solving framework that solves a newly encountered problem (i.e. meal insulin dosing), based on the information obtained from previously solved problems (cases). CBR is usually described in four steps: *Retrieve* the most similar cases from a case-base (e.g. late dinner preceded by moderate exercise); *Reuse* solutions of retrieved cases (e.g. bolus calculator parameters ICR and ISF); *Revise* the outcome of the applied solution (e.g. post-prandial glucose excursion); and *Retain* the new cases if considered useful for solving future problems [1]. In ABC4D, cases are stored in a case-base representing meal scenarios with significantly different insulin requirements (e.g. breakfast after exercise vs. dinner after watching a movie) and therefore, requiring a different insulin dosing. Retrieving of the cases was performed by means of an Euclidian distance with equal weights on all parameters. It is important to note that, unlike the traditional CBR approach where solutions of cases in the case-base are static, in ABC4D such a solutions (i.e. ICR and ISF) are adapted if considered to be sub-optimal. In order to perform such adaptation of sub-optimal solutions, a modified version of Run-to-Run algorithm proposed by Herrero et al. [4] is employed.

2.3 Run-to-Run Control (R2R)

R2R is a control methodology designed to exploit repetitiveness in the process that is being controlled [8]. Its purpose is to enhance performance, using a mechanism of trial and error. Owens et al. [9] used this idea to exploit the repetitive nature of the insulin therapy regimen of the diabetic patient. However, the requirement of one pre-prandial capillary blood glucose measurement and two post-prandial ones made the approach impractical. The simplest formulation of R2R may be,

$$u_{k+1} = u_k + K \cdot error, \quad (2)$$

where u is the control action, K is a tuning gain and error is the tracking error defined as the difference between a measurement from the process and a set-point. The R2R algorithm used in ABC4D is based on the hypothesis that the meal insulin bolus can be adjusted based on the residual between the minimal post-prandial glucose concentration (G_{min}) obtained with a continuous glucose monitor (CGM) and a predefined glucose set-point (G_{sp}) over a predefined time window $[t_1, t_2]$. Therefore, the updated bolus is calculated as

$$B_{k+1} = B_k + K \cdot (G_{min} - G_{sp}), \quad (3)$$

where $K \cdot (G_{min} - G_{sp})$ is the extra insulin that needs to be added (or subtracted) to the original bolus (B_k) in order to bring blood glucose levels back to the set-point (G_{sp}), and K is defined as $K = 1/ISF$. In order to provide robustness to the metric against the inherent variability and uncertainty of the system (e.g. sensor noise and carbohydrate estimation), a glucose range $[G_l, G_h]$ is defined where no adaptation is done if G_{min} falls within this range.

However, the ABC4D R2R algorithm is not fully suited to be used within a closed-loop (CL) controller. Note that the CL controller can compensate for the lack of meal bolus and still bring glucose levels within the target range $[G_l, G_h]$, but the post-prandial glucose peak can still be significantly sub-optimal. Assuming that the CL controller is correctly tuned, the ABC4D R2R metric is still valid when G_{min} falls below the target range. Otherwise, a new metric for adjusting ICR is required. The new proposed metric is based on the hypothesis that, assuming that the CL controller is appropriately tuned, the insulin delivered by the CL controller during the postprandial period over the basal insulin, is insulin that should have been delivered by the meal-priming bolus. Thus, the bolus calculator parameters can be updated based on this additional insulin. Therefore, Equation 3 is replaced by

$$if \ G \leq G_l \ B_{k+1} = B_k + K \cdot (G_{min} - G_{sp}), \quad (4)$$

$$else \ B_{k+1} = B_k + \sum_{t_3}^{t_4} D(t), \quad (5)$$

where $D(t)$ is the insulin delivered by the controller over the basal insulin level during the time window $[t_3, t_4]$ and glucose levels are over G_h .

Assuming the correlation $ISF = (1960 \cdot ICR) / (2.6 \cdot W)$ reported by Walsh et al. [7], where W is the subjects weight (lbs), the updated ICR can be calculated from Equation 1 as

$$ICR_{k+1} = \frac{CHO + \frac{(G_{min} - G_{sp})}{1960/(2.6 \cdot W)}}{B_{k+1} + IOB}. \quad (6)$$

2.4 In Silico Evaluation

The latest version of the UVa-Padova T1DM simulator (v3.2) (Epsilon Group, MA, US) was used to evaluate the proposed adaptive bolus calculator for closed-loop controllers. The 11 adult subjects available in the simulator were used for this purpose. A three-month scenario was selected in order to leave enough time to the meal bolus adaptation mechanism to converge. Inter- and intra-subject variability of insulin requirements and uncertainty on carbohydrate intake were considered as proposed by Herrero et al [4]. It is important to remark that due to the inherent limitations of the simulator, only three cases (i.e. breakfast, lunch and dinner) were considered by the CBR algorithm. Nevertheless, initial clinical trials of the ABC4D algorithm show promising results [10]. The following standard glycemic control metrics were selected for comparison purposes: mean blood glucose (BG); percentage time in target range $[70, 180]$ mg/dl ($\%inT$); percentage time below target ($\% < T$); percentage time above target ($\% > T$); and daily average of insulin delivered in units of insulin (TDI).

3 Results

Table 1 shows the results corresponding to the 11 adults for each one of the evaluated control strategies (AP vs. ABC-AP).

4 Conclusion

Integrating an adaptive meal bolus calculator within the Imperial College Artificial Pancreas controller significantly improves all the evaluated glycemic outcomes in a virtual type 1 diabetes population (11 adults) when compared against the Imperial College Artificial Pancreas without bolus adaptation over a three-month scenario with realistic inter-subject and intra-day variability. It is worth noting that the significant reduction in hyperglycemia was achieved without an any increase in hypoglycemia. Trials have been planned to clinically validate the proposed technique.

REFERENCES

- [1] A Aamodt and E Plaza, 'Case-based reasoning: Foundational issues, methodological variations, and system approaches', *AI communications*, **7**(1), 39–59, (1994).
- [2] Herrero et al., 'Advanced insulin bolus advisor based on run-to-run control and case-based reasoning', *Biomedical and Health Informatics, IEEE Journal of*, **19**(3), 1087–1096, (2015).
- [3] El-Khatib et al., 'Autonomous and continuous adaptation of a bihormonal bionic pancreas in adults and adolescents with type 1 diabetes.', *A J Clin Endocrinol Metab.*, **99**(5), 1701–11, (2014).
- [4] Herrero et al., 'Method for automatic adjustment of an insulin bolus calculator: in silico robustness evaluation under intra-day variability.', *Comput Methods Programs Biomed.*, **119**(1), 1–8, (2015).
- [5] Reddy et al., 'Metabolic control with the bio-inspired artificial pancreas in adults with type 1 diabetes: A 24-hour randomized controlled crossover study.', *J. Diabetes Sci. Technol.*, **10**(2), 1405–13, (2015).
- [6] Thabit H et al., 'Home use of an artificial beta cell in type 1 diabetes.', *N Engl J Med*, **373**, 2129–2140, (2015).
- [7] Walsh et al., 'Guidelines for optimal bolus calculator settings in adults', *J. Diabetes Sci. Technol.*, **5**(1), 129–135, (2011).
- [8] Wang et al., 'Survey on iterative learning control, repetitive control, and run-to-run control.', *Journal of Process Control.*, **19**(10), 1589–1600, (2009).
- [9] C. Owens, H. Zisser, L. Jovanovic, B. Srinivasan, D. Bonvin, and Doyle FJ 3rd., 'Run-to-run control of blood glucose concentrations for people with type 1 diabetes mellitus', *IEEE Trans. Biomed. Eng.*, **53**, 996–1005, (2006).
- [10] M. et al. Reddy, 'Clinical safety and feasibility of the advanced bolus calculator for type 1 diabetes based on case-based reasoning: a 6-month randomised single-arm pilot study.', *Diabetes Technol Ther.*, **Epub ahead of print**, (2016).

Table 1. Glycemic results corresponding to the 11 adult subjects.

	BG	% <i>inT</i>	% < <i>T</i>	% > <i>T</i>	TDI
AP	142.2 ± 9.4	82.0 ± 7.0	0.21 ± 0.36	17.7 ± 7.0	45.8 ± 10.1
ABC-AP	131.8 ± 4.2	89.5 ± 4.2	0.21 ± 0.18	10.2 ± 4.1	48.5 ± 10.4
p	< 0.001	< 0.001	0.99	< 0.001	0.002

Temporal case-based reasoning for bolus decision support

Daniel Brown, Rachel Harrison, Clare Martin and Ian Bayley¹

Abstract. Individuals with type 1 diabetes frequently have to determine what quantity of bolus insulin is required at meal time in order to maintain their blood glucose levels. To help this process bolus calculators have been developed to suggest appropriate doses. However, these calculators do not automatically adapt to improve bolus suggestions and instead require fine tuning of certain parameters, a process that often requires clinical input.

To overcome these limitations, we suggest using the artificial intelligence technique case-based reasoning to personalise bolus decision support. A novel aspect of our approach is the use of temporal sequences to factor in preceding events to the decision making process as opposed to looking at events in isolation.

The *in silico* results of the approach show that the temporal retrieval algorithm successfully identifies appropriate cases for reuse. Additionally through insulin-on-board adaptation and postprandial revision, the approach is able to learn and improve bolus predictions, reducing the blood glucose risk index by up to 27% after three revisions of a bolus solution.

1 INTRODUCTION

Type 1 diabetes mellitus (T1DM) is a condition is caused by a defective autoimmune system, leading to the destruction of pancreatic beta cells. This results in an individual's inability to automatically control their blood glucose levels. To overcome this the individuals must carefully manage their condition to avoid *hypoglycaemia* (low blood glucose levels) and *hyperglycaemia* (high blood glucose levels), both of which can have serious health implications.

Bolus insulin calculators are available to assist management of the condition, which are shown to be effective [2]. However, these bolus calculators will always produce the same result from the user's inputs unless certain settings such as the carbohydrate-to-insulin ratio (CIR) and insulin sensitivity factor (ISF) are altered, a process often guided by clinicians; where the CIR is the number of carbohydrates covered by a unit of insulin, whilst the ISF is the drop in blood glucose per unit of insulin. It is this problem our research aims to address through replacing the static formula with the ability to learn and improve bolus recommendations automatically through case-based reasoning (CBR).

We begin by briefly explaining the fundamentals of CBR in Section 2, highlighting the limitation of using cases in isolation in temporal domains such as T1DM. In Section 3 we describe our approach to solving this problem using CBR. Section 4 outlines the results of this approach, showing the system's ability to improve results over time. We then discuss related work in section 5. Finally, conclusions reached are described in Section 6.

2 CASE-BASED REASONING

Case-based reasoning is a well-established form of artificial intelligence which attempts to mimic the human ability to recall appropriate solutions to problems. The foundations of CBR can be found in the pioneering work conducted by Kolodner based on the idea of dynamic memory modelling proposed by Schank [13, 19].

A widely adopted CBR model is the R^4 model proposed by Aamodt and Plaza [1]. The R^4 model is four stage cycle: retrieve, reuse, revise, and retain. Firstly, a new problem is presented to the system. Based on the features and feature-values of the problem, a similar case is retrieved. The retrieved case is then reused to solve the new problem; this may involve some form of adaptation to resolve any discrepancies between the proposed problem and the retrieved case. A solution is then presented, which subject to real-world or simulated use can be further revised. Once the solution is accepted it is retained in the case-base. This cycle then continues, with each new problem having a larger and/or refined case-base to aid predicting solutions of future problems.

The majority of research and development using CBR considers each case to an isolated event. In the context of T1DM we believe that temporal effects should be factored into the retrieval step. Research into temporal CBR has been relatively limited, with the majority of methods requiring specialist case representation, e.g. [11, 12]. To overcome this, sequences of continuous temporal cases can be merged into a singular case [18]. This method allows the temporal sequences to be compared using standard distance metrics without the need for additional rules. Plausible episodes are generated from a new problem, which are then compared to similar retrieved episodes in order to solve the new problem. We use this formation of episodes as the foundation for our temporal approach.

3 TEMPORAL CASE-BASED REASONING FOR BOLUS INSULIN DECISION SUPPORT

This section discusses our approach to using the R^4 model in the context of bolus advice [6, 5]. We begin by defining the structure of cases, then describe each step of the R^4 model.

3.1 Case structure

Unlike other CBR systems where case features may vary, in this context the features representing a case are well-defined. The initial step taken by this research was to determine which parameters are required by bolus calculators. Through assessment of existing bolus calculators it was found that the parameters described in Table 1 are used by the Accu-Chek Aviva Expert (AE), RapidCalc (RC), Diabetes Personal Calculator (DPC), Diabetic Dosage (DD), and InsulinCalc (IC). The apps were selected using a method described by Martin et. al. [16].

¹ Oxford Brookes University, email: [dbrown, rachel.harrison, cemartin, ibayley]@brookes.ac.uk

Table 1. Parameters used by existing bolus calculators

Parameter	AE	RC	DPC	DD	IC
Carbohydrate intake	✓	✓	✓	✓	✓
Preprandial blood glucose	✓	✓	✓	✓	✓
Target blood glucose	✓	✓	✓	✓	✓
Insulin sensitivity factor	✓	✓	✓	✓	✓
Carbohydrate-to-insulin ratio	✓	✓	✓		✓
Insulin-on-board	✓	✓	✓		
Exercise	✓	✓			

The parameters identified from the existing bolus calculators in Table 1 allow us to describe the features of a case. It is clear that the carbohydrate intake, preprandial blood glucose level, and target blood glucose level are essential case parameters.

The ISF and CIR are the primary parameters used to tune the bolus calculator. These will be omitted from the cases since the CBR approach seeks to replace their role in the decision making process. Instead they will be replaced by the date and time of the event, since they are usually defined as personal settings on the device and largely remain static, making them redundant to the CBR retrieval step.

Insulin-on-board (IOB) is a crucial parameter which helps to avoid the negative effects of insulin stacking, caused by administering insulin when some already remains active in the body. To cater for IOB, the *retrieve* step (Section 3.2) uses a temporal approach that factors in preceding bolus doses. This is coupled with an adaptation rule in the *reuse* step (Section 3.3), which resolves differences between the IOB in the problem and the retrieved case(s).

Exercise is a parameter that we believe should be included. However, the UVa/Padova T1DM simulator [14] used in this research did not allow this to be modelled, so it must be omitted.

Finally, the *solution* needs to be retained by the case for reuse in solving new problems. The solution is this approach is the bolus dose. This will also serve as a feature in temporal aspect described in the *retrieve* step.

Following the assessment of parameters used by bolus calculators, we decide that cases will be represented by the date and time, carbohydrate intake, preprandial blood glucose level, and the solution of bolus.

3.2 Retrieve

The retrieval step is where the temporal aspect is introduced to the system. As opposed to looking at the new problem and previous cases in isolation, we believe the bigger picture should be considered, most notably preceding events. Whilst the temporal side of CBR has been considered previously, none of the previous methods appear suitable for the task of bolus decision support. To address this, we propose the use of a *temporal sequence* to describe both new problems and previous cases based upon a method described by [18].

Definitions 3.1 through to 3.4 describe the method more formally. In Def. 3.1 and Def. 3.2 a case and the case-base for an individual patient are defined.

Definition 3.1 (Case) A case c is a tuple comprised of a number of n features f_i , together with a solution s .

$$c = (f_1, f_2, \dots, f_n, s)$$

Definition 3.2 (Case-base) A case-base CB is a sequence of cases c_i , where i ranges from 1 to the size of the case-base.

$$CB = \langle c_1, c_2, \dots, c_{|CB|} \rangle$$

The temporal sequence describing the new problem TP is defined in Def. 3.3. A TP with $t = 1$ will be a sequence containing the new problem c' , resulting in traditional CBR retrieval where no previous events are included. For a TP with $t > 1$, the sequence must start from $t - 2$ less than the size of the case-base, because at the very least the sequence must contain the new problem c' and the last case in the case-base $c_{|CB|}$.

Definition 3.3 (Temporal problem sequence) A temporal problem sequence TP is comprised of the individual new problem proposed to the system c' together the preceding cases c in the case-base ordered by date and time. The size of TP is determined by the defined temporal sequence length t , where $1 \leq t \leq |CB|$.

$$TP = \langle c_{|CB|-(t-2)}, c_{|CB|-(t-3)}, \dots, c_{|CB|}, c' \rangle$$

The problem sequence is then compared to sequences in the case-base (Def. 3.4) of the same temporal sequence length t . The sequences must be the same length in order to conduct similarity, a process that will identify the most relevant sequence in the case-base.

Definition 3.4 (Temporal case sequence) A temporal case sequence TC_n is comprised of the case c_n together with $t - 1$ preceding cases ordered by date and time, where t is the sequence length.

$$TC_n = \langle c_{n-(t-1)}, c_{n-(t-2)}, \dots, c_n \rangle$$

To deal with broken sequences - those with assumed missing events (gaps) - the outer fence defined by Tukey is used [20]. Where such gaps exist, the features are replaced by the maximum distance of 1 on the scale [0, 1].

A weighted distance function is used to compare the similarity of TP and TC_n , this helps to ensure that the importance of each feature on the overall similarity is representative of the problem. Feature weightings were determined using the Weka data mining tool, which includes the feature selection algorithms: Chi-Squared, Information Gain, Gain Ratio, One Rule, RELIEF-F and Symmetrical Uncertainty [21]. All the aforementioned feature selection algorithms are single-attribute evaluators and return a score determining each attribute's likelihood to predict the class (bolus dose). To derive the feature weightings sample data sets were produced using closed-loop simulation [14]. Cases were then extracted from the simulation output, merged into single cases representing temporal sequences, and finally processed using Weka.

The weighted Euclidean distance function for determining similarity is described in Eq. 1. Let TP and TC be the problem and case sequences respectively, I be the total number of features, and w be the weight of the respective feature. Prior to computing the distance, all features are normalised to avoid unwanted bias.

$$d(TP, TC) = \sqrt{\sum_{i=1}^I w_i (TP_i - TC_i)^2} \quad (1)$$

3.3 Reuse

For the reuse step we adopted a simple k -NN regression strategy to average the bolus prediction of k retrieved cases. Equation 2 defines the reuse strategy, let k define the number of retrieved case, and i_n define bolus solution provided by a retrieved case.

$$\text{suggested bolus dose} = \frac{1}{k} \sum_{n=1}^k i_n \quad (2)$$

The result is then adapted to resolve differences in the IOB from the new problem to the retrieved cases to further tune in the bolus recommendation. Whilst the use of temporal sequences somewhat resolves this issue alone, it is important to prevent the negative effects of insulin stacking. In this research a linear IOB algorithm (Eq. 3) is adopted [7].

The adapted bolus suggestion is calculated by deducting the average of the sum of the IOBs for all the retrieved cases from the original bolus suggestion to determine the difference d' , as described in Eq. 6. For Eq. 3 - 6 the variables are defined as follows: the case-base CB is a sequence of cases c , with each case c a tuple of case time ct in minutes and the bolus dose ci . t denotes time in minutes, pt is the time of a new problem in minutes. RC denotes a sequence of case times in minutes. The active insulin time a is a constant to reflect the duration of a bolus dose in minutes. The suggested bolus dose i is the original bolus dose to be adapted.

$$\text{iob}(c, t, a) = \begin{cases} ci \times \left(1 - \frac{t - ct}{a}\right), & \text{if } a > t - ct > 0 \\ 0.0, & \text{otherwise.} \end{cases} \quad (3)$$

$$\text{iob}_s(CB, t, a) = \sum_{n=1}^{|CB|} \text{iob}(c_n, t, a) \quad (4)$$

$$d(pt, RC, CB, a) = \text{iob}_s(CB, pt, a) - \frac{\sum_{n=1}^k \text{iob}_s(CB, RC_n, a)}{k} \quad (5)$$

$$d' = \begin{cases} i - d(pt, RC, CB, a), & \text{if } i - d(pt, RC, CB, a) \geq 0 \\ 0.0, & \text{otherwise.} \end{cases} \quad (6)$$

3.4 Revise

The revise step is crucial to allow the system to improve sub-optimal recommendations. The degree of success can be inferred from the difference between postprandial blood glucose of the subject and their target blood glucose level. If the postprandial reading is equal or close to the target blood glucose level then the recommendation can be considered optimal and no revision is required. However, if the postprandial reading is higher or lower than the target level, the recommended bolus should be increased or decreased respectively.

To determine this, a method for correcting bolus doses described by Eq. 8 is used based on the subject's total daily dose to estimate the ISF (Eq. 7) [3, 9]. Let I represent the sequence of bolus and basal doses over a period of d days, I_i be an individual bolus or basal dose from the sequence of insulin doses I , pbg be a postprandial blood glucose reading (mmol/L), and tbg be the target blood glucose level (mmol/L).

$$ISF = (1700 \div \sum_{i=1}^{|I|} I_i) \times 0.0555 \text{ mmol/L} \quad (7)$$

$$\text{revised bolus} = \frac{pbg - tbg}{ISF} \quad (8)$$

One difficulty to overcome is when to perform the postprandial blood glucose reading. If it occurs too soon after the dose was administered or too late then the revision is likely to be sub-optimal. To determine this, *in silico* results for 2, 3 and 4-hour offsets were evaluated, with 3-hour found to be the most optimal.

3.5 Retain

The retain step of the cycle stores the evaluated recommendation into the case-base for future reuse. The complexity of retaining cases largely depends on how the cases are stored. In this work we did not place much emphasis on this step since the case structure remains consistent. However, we are aware of the importance of case-base maintenance to ensure the search space does not cause time-complexity issues, and to prevent bad solutions being retained.

4 RESULTS

In this section we describe the *in silico* results of the approach outlined in Section 3. The results are broken down into the first three steps of the CBR cycle to highlight how these different steps (retrieve, reuse, and revise) of our approach help to progressively improve the decisions made by the system.

4.1 Retrieve

Five sets of new problems were created to test against the case-bases. The problem sets contained one month of new problems (approximately 130-140 problems), allowing us to observe the improvements in blood glucose prediction from the solutions obtained during retrieval. Each of the problem sets was applied to each of the case-bases for 1 to 5 nearest neighbours with six different single-attribute feature evaluators (Chi-Squared, Information Gain, Gain Ratio, One Rule, RELIEF-F, and Symmetrical Uncertainty) [21].

The blood glucose risk index (BGRI) was the primary statistical measure we used to measure our predictions. This measure can be applied to continuous blood glucose data to determine overall variance of a low blood glucose risk index (LBGI) and high blood glucose risk index (HBGI) [8].

Table 2 presents the percentage change in BGRI of the different temporal sequence lengths (TS2 - TS5) in comparison to no temporal sequence (TS1) for all feature selection algorithms, where the highest percentage reduction in BGRI result is best. This result illustrates that temporal sequences provide some improvement in case retrieval.

Table 2. Percentage change in BGRI for different temporal sequence lengths (TS2 - TS5) to no temporal sequence (TS1)

Feature selection algorithm	TS1 BGRI	TS2 %	TS3 %	TS4 %	TS5 %
Chi-Squared	4.44	-1.07	-1.02	-0.58	-0.29
Information Gain	4.43	-1.21	-0.95	-0.74	-0.23
Gain Ratio	4.44	-1.26	-1.11	-0.83	-0.49
One Rule	4.42	-0.52	-0.60	-0.81	-1.23
RELIEF-F	4.43	-0.72	-0.12	-0.26	-0.26
Symmetrical Uncert.	4.43	-1.20	-1.03	-0.84	-0.39

4.2 Reuse

Insulin-on-board adaptation was tested against a combination of five case-base sets using the optimal retrieval configuration. The purpose of the IOB adaptation is to resolve the differences in active insulin between the new problem and retrieved case(s). Table 3 illustrates the improvement the IOB adaptation provides across all statistical measures.

Table 3. Comparison without and with insulin-on-board adaptation

Measure	Without IOB	With IOB
BGRI	4.22 \pm 0.31	3.94 \pm 0.27
< target range (TR) %	0.03 \pm 0.19	0.01 \pm 0.12
> target range (TR) %	0.00 \pm 0.00	0.00 \pm 0.00
σ	0.87 \pm 0.05	0.81 \pm 0.04
μ mmol/L	6.34 \pm 0.13	6.30 \pm 0.21

4.3 Revise

As stated previously, successful revision is crucial for CBR to learn from mistakes. Table 4 presents the *in silico* results of one to three cycles of 3-hour offset postprandial revision, where the *original* bolus is after reuse adaptation, but prior to revision. The results demonstrate how the postprandial revision rule improves suggestions based using the difference between a target blood glucose level and a postprandial blood glucose reading. After three revisions the resulting BGRI is reduced by as much as 27% from the original bolus suggestion.

Table 4. Bolus reuse following 3-hour offset postprandial evaluation

Measure	Original	Cycle 1	Cycle 2	Cycle 3
BGRI	3.94 \pm 0.27	3.32 \pm 0.31	3.02 \pm 0.41	2.87 \pm 0.43
< TR %	0.01 \pm 0.12	0.00 \pm 0.00	0.00 \pm 0.00	0.00 \pm 0.00
> TR %	0.00 \pm 0.00	0.00 \pm 0.00	0.00 \pm 0.00	0.00 \pm 0.00
σ	0.81 \pm 0.04	0.72 \pm 0.03	0.67 \pm 0.04	0.65 \pm 0.04
μ mmol/L	6.30 \pm 0.21	6.44 \pm 0.15	6.52 \pm 0.13	6.56 \pm 0.12

5 RELATED WORK

Case-based reasoning has been adopted by several research projectors in the domain of T1DM. The majority of this research has focused on aiding clinicians with therapy adjustments as opposed to the patient directly. Such projects include the T-IDDM project [4], and more recently the IDSDM project [15]. A notable exception is the Advanced Bolus Calculator for Diabetes (ABC4D) [10, 17], which through clinical trials demonstrated the positive affects of CBR for bolus advice. Whilst ABC4D tackles the same problem, we adopt a different approach to CBR, and incorporate the temporal aspect.

6 CONCLUSION

This research demonstrated positive *in silico* results for the use of temporal CBR for bolus decision support. The introduction of a temporal retrieval algorithm demonstrated an improved BGRI prior to any adaptation of revision. With the introduction on IOB adaptation and a postprandial revision algorithm, a notable improvement in all statistical measures is demonstrated. These results highlight the potential benefit of temporal CBR for bolus decision support over bolus calculators currently available to the public.

We are aware of limiting factors in this research, most notably the inability to include additional factors such as physical exercise due to limitations of the simulator. Further research of this approach should include additional parameters, a safety layer to protect patients, and validation through clinical trials.

ACKNOWLEDGEMENTS

We thank Oxford Brookes University for funding this research.

REFERENCES

- [1] A. Aamodt and E. Plaza, 'Case-based reasoning: Foundational issues, methodological variations, and system approaches', *AI Communications*, **7**, 39–59, (1994).
- [2] K. Barnard, C. Parkin, A. Young, and M. Ashraf, 'Use of an automated bolus calculator reduces fear of hypoglycemia and improves confidence in dosage accuracy in patients with type 1 diabetes mellitus treated with multiple daily insulin injections', *Journal of Diabetes Science and Technology*, **6**(1), 144–149, (2012).
- [3] Becton, Dickinson and Company, *Staying on Target™: Your Insulin Adjustment Workbook. Yes, You Can Do It!*, Becton, Dickinson and Company, 2005.
- [4] R. Bellazzi, C. Larizza, S. Montani, A. Riva, M. Stefanelli, G. d'Annunzio, R. Lorini, E. Gómez, et al., 'A telemedicine support for diabetes management: the T-IDDM project', *Computer methods and programs in biomedicine*, **69**(2), 147–161, (2002).
- [5] D. Brown, *Temporal Case-based Reasoning for Insulin Decision Support*, Ph.D. dissertation, Department of Computing and Communication Technologies, Oxford Brookes University, 2015.
- [6] D. Brown, I. Bayley, R. Harrison, and C. Martin, 'Developing a mobile case-based reasoning application to assist type 1 diabetes management', in *2013 IEEE 15th International Conference on e-Health Networking, Applications & Services (Healthcom)*, Lisbon, Portugal, (2013). IEEE.
- [7] R. Campbell and A. Abramovich, 'Calculating insulin on board for extended bolus being delivered by an insulin delivery device', 2012. US Patent 8,140,275.
- [8] W. Clarke and B. Kovatchev, 'Statistical tools to analyse continuous glucose monitor data', *Diabetes Technology and Therapeutics*, **11**, S–45–S–55, (2009).
- [9] P. Davidson, H. Hebblewhite, B. Bode, R. Steed, N. Welch, M. Greenlee, P. Richardson, and J. Johnson, 'Statistically based CSII parameters: correction factor, CF (1700 rule), carbohydrate-to-insulin ratio, CIR (2.8 rule), and basal-to-total ratio', *Diabetes Technology and Therapeutics*, **5**(5), A237, (2003).
- [10] P. Herrero, P. Pesl, M. Reddy, N. Oliver, P. Georgiou, and C. Toumazou, 'Advanced insulin bolus advisor based on run-to-run control and case-based reasoning', *IEEE Journal of Biomedical and Health Informatics*, **19**, 1087–1096, (2015).
- [11] M. Jaczynski, 'A framework for the management of past experiences with time-extended situations', in *Proceedings of the sixth international conference on Information and knowledge management*, pp. 32–39, Las Vegas, NV, USA, (1997). ACM.
- [12] M. Jære, A. Aamodt, and P. Skalle, 'Representing temporal knowledge for case-based prediction', in *Advances in case-based reasoning*, 174–188, Springer, (2002).
- [13] J. Kolodner, *Case-based Reasoning*, Morgan Kaufmann, 1993.
- [14] B. Kovatchev, M. Breton, C. Dalla Man, and C. Cobelli, 'In silico pre-clinical trials: a proof of concept in closed-loop control of type 1 diabetes', *Journal of Diabetes Science and Technology*, **3**, 44–55, (2009 2009).
- [15] C. Marling, J. Shubrook, and F. Schwartz, 'Case-based decision support for patients with type 1 diabetes on insulin pump therapy', *Advances in Case-Based Reasoning, Lecture Notes in Computer Science*, **5239**, 325–339, (2008).
- [16] C. Martin, D. Flood, D. Sutton, A. Aldea, R. Harrison, and M. Waite, 'A systematic evaluation of mobile applications for diabetes management', in *INTERACT 2011, 13th IFIP TC 13 International Conference, Proceedings, Part IV, Human-Computer Interaction*, pp. 466–469, Lisbon, Portugal, (2011). Springer.
- [17] P. Pesl, P. Herrero, M. Reddy, N. Oliver, D. Johnston, C. Toumazou, and P. Georgiou, 'Case-based reasoning for insulin bolus advice: Evaluation of case parameters in a six-week pilot study', *Journal of Diabetes Science and Technology*, (2016).
- [18] M. Sánchez-Marré, U. Cortés, M. Martínez, J. Comas, and I. Rodríguez-Roda, 'An approach for temporal case-based reasoning: Episode-based reasoning', in *Case-Based Reasoning Research and Development*, 465–476, Springer, (2005).
- [19] R. Schank, *Dynamic Memory: A Theory of Reminding and Learning in Computers and People*, Cambridge University Press, 1983.
- [20] J. Tukey, *Exploratory data analysis*, Pearson, 1977.
- [21] I. Witten and E. Frank, *Data Mining: Practical Machine Learning Tools and Techniques, Second Edition (Morgan Kaufmann Series in Data Management Systems)*, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2005.

Monitoring patients with diabetes using wearable sensors: Predicting glycaemias using ECG and respiration rate

Božidara Cvetković¹ and Urška Pangerc¹ and Anton Gradišek¹ and Mitja Luštrek¹

Abstract. Wearable sensors show great promise in monitoring medical conditions of patients with diabetes and can therefore be used to significantly improve their quality of life. In our pilot study, patients with type I and II diabetes were equipped with a series of such sensors. Here, we focus on the data provided by a chest harness sensor that records both the ECG signal and the respiration rate. We developed machine-learning based models to recognise and predict abnormal glucose blood levels (hypo- and hyperglycaemia) in type I and II diabetes patients. We obtained 84 % accuracy in predicting glycaemia for patients with type I diabetes and 88 % for patients with type II. For recognition of glycaemia, we achieved 78 % accuracy for type I and 76 % for type II. Analysis of other sensor data is in progress.

1 INTRODUCTION

Diabetes is a group of chronic metabolic diseases that are related to high blood sugar (glucose) levels, either due to the pancreas not producing enough insulin or the body not properly responding to it. The two main types are type I, which is an autoimmune condition where the immune system destroys insulin-producing cells in the pancreas, whereas type II is related to insulin resistance and is primarily caused by unhealthy living style. According to the International Diabetes Federation, diabetes currently affects over 400 million people worldwide (out of which, 90% is type II), reaching epidemic proportions, with numbers expected to rise up to 600 million in 20 years [1].

People with diabetes have to adjust their lifestyle in order to keep the blood sugar in the appropriate range, in order to prevent medical complications that may otherwise arise – especially the cardiovascular diseases, stroke, chronic kidney failure, damage to the eyes or food ulcers. Diabetes-related complications also represent the 8th leading cause of death worldwide.

In the last decade, wearable sensors for a variety of purposes have become widely available. They can be used to track basic body functions, such as the respiration rate, ECG, body temperature, or even more complex features such as types of activities and energy expenditure [2] through an efficient interpretation of accelerometer, gyroscope, or other available biosensor data [3]. A combination of different types of information can assist individual patients in monitoring their medical condition, such as predicting the blood glucose levels and early warning of (preventable) potential complications, thus greatly improving their quality of life. In a pilot study, carried out in the framework of the

COMMODITY12 EU project [4], a group of patients was equipped with a series of sensors, wearable and non-wearable, in order to assess feasibility and extent to which these sensors can assist patients in everyday life.

In this paper, we focus on the interpretation of the ECG and the respiration rate data (obtained using a commercial chest harness sensor [5]) in combination with continuous blood glucose level measurements obtained with GlucoTel [6], a telemedical blood glucose measuring sensor. These data were used for development of two machine-learning based models, one for detection of potential hypo- and hyperglycaemias and one for predicting their occurrences. We discuss potential improvements in combination with data from other sensors as well as in combination with more complex features which already utilise machine-learning (e.g., recognised activities, estimated energy expenditure, etc.).

2 RELATED WORK

Medical literature states that hypoglycaemia (low glucose levels) is related to decrease in heart rate and that hyperglycaemia (high levels) is strongly linked to the polarisation and depolarisation of heart chambers, the so-called QT interval in the electrocardiogram readings (discussed in Section 3.1). These changes in the QT interval are also highly linked to arrhythmias which can lead to cardiac arrest or heart failure.

Hanfeld et al. [7] present a systematic overview of the state-of-the-art in the field. For patients with type I and II diabetes, it was found that the changes of the QT interval occur in cases of severe hypoglycaemia. Other studies [8,9] also reached the same conclusion. On the other hand, Singh et al. [10] demonstrated that the heart rate variability decreases in case of severe hyperglycaemia.

Nguyen et al. [11] attempted to detect hypo- and hyperglycaemias from the ECG signal from patients with type I diabetes. They found that an increasing heart rate relates exclusively to hypoglycaemia while changes of the PR interval from ECG exclusively relate to hyperglycaemia.

Machine-learning algorithms have previously been used to predict the blood glucose levels [12]. However, these algorithms use complex dynamic models based on historic data for individual patient as the input parameters, and not the ECG measurements.

In the related research, the ECG signal was typically measured with professional equipment under clinical supervision in the laboratory environment. The researchers could immediately discard the noisy data and therefore investigated the correlation between the values of the ECG parameters and the measured glucose levels only on clean data. Our research motivation is to detect and predict

¹ Department for Intelligent Systems, Jožef Stefan Institute, Jamova 39, SI-1000 Ljubljana, Slovenia

hyper- and hypoglycaemias in everyday lives of patients and not only under medical supervision. That is why the pilot study utilised a commercial ECG sensor [5] that the patients have worn at home.

3 DATASET

The dataset was collected during the project pilots in two countries. The study encompassed 30 patients with type II diabetes from Poland and 22 patients with type I diabetes from Italy. Each patient was equipped with Zephyr BioHarness [5] that records ECG, respiration rate, and acceleration, with a GlucoTel [6] glucose monitor, with a standard telemetric blood pressure monitor, a telemetric scale, and a smartphone that was used as a smart-hub which serves as a main control system which enables input of symptoms (e.g., tremor, vertigo, etc.), collects the data from all devices and sends the data to the central server.

The patients were instructed to wear the ECG sensor and perform measurements while performing normal daily activities (eating, exercising) and around the time they measured glucose level, over a course of six weeks. In total, we have collected 787 hours of raw ECG and respiration data during the pilot study.

In the study at hand, we analysed the ECG signal, the respiration rate, and the glucose level measurements. To obtain clean data we first processed the ECG and the respiration rate measurements using filters which removed the noisy and unreadable parts, but nevertheless retained the signal morphology. After filtering, we were left with approximately 566 hours of clean ECG and respiration data.

With respect to the glucose level measurements, two types of 30-minutes segments were used for analysis:

- 30 minute segment from 45 to 15 minutes before the glucose measurement, for the purpose of glucose level prediction (hypo- and hyperglycaemia and normal levels)
- 30 minute segment from 15 minutes before to 15 minutes after the glucose measurement, for the purpose of glucose level detection

The **ECG signal** was processed with an ECG feature extraction algorithm [13] that extracts 13 parameters which describe the shape of the signal (Figure 1).

The parameters are the following:

- PR segment – time between the end of the P wave and beginning of the QRS complex
- PR interval – time between the beginning of the P wave and the beginning of the QRS complex
- QS interval – time between the beginning and the end of the QRS complex
- ST segment – time between the end of the QRS complex and beginning of the T wave
- QT interval – time between the beginning of the QRS complex and end of the T wave
- P wave length – time between the beginning and the end of the P wave
- T wave length – time between the beginning and the end of the T wave
- Q, R, S, P, and T values – the amplitudes of the Q, R, S, P, and T waves, respectively (as individual parameters)

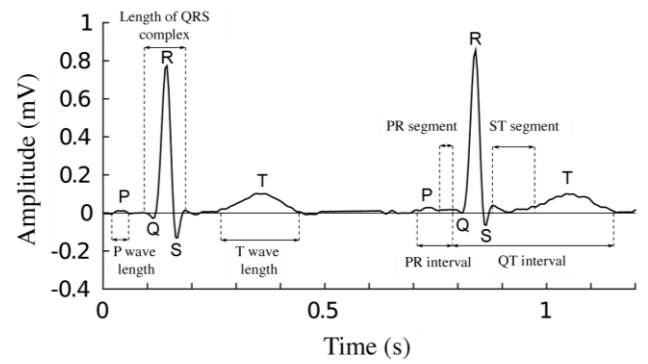


Figure 1. ECG parameters retrieved with the signal processing algorithm.

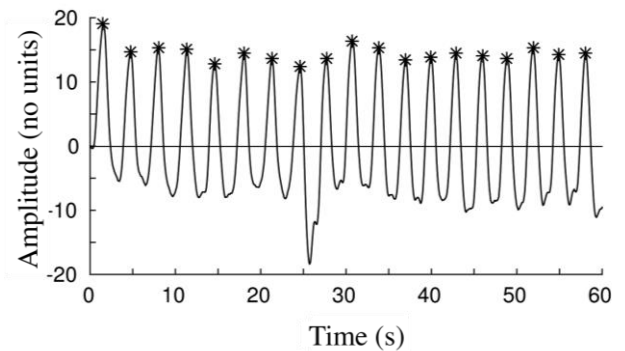


Figure 2. Respiration signal. The stars are labels which indicate the recognised breaths.

- RR interval – time between two consecutive R waves

For each of the parameters, the average value, the standard deviation, and the trend (the slope of a linear approximation) were calculated over the whole 30-minute interval.

The signal of the **respiration rate** measurements is shown in Figure 2. As with the ECG signal, the average respiration rate was calculated, together with the standard deviation and the trend.

The **glucose level measurements** contain, apart from the glucose level value itself, also the information whether the measurement was done before or after a meal, before sleep, at night, or other. Patients were left to decide when to take the measurements. Based on the glucose levels, the measurements were sorted into three groups as presented in Table 1.

Table 1. Type of glycaemia according to the measured glucose level.

Glycaemia	Glucose level
Hypoglycaemia	< 4 mmol/l
Hyperglycaemia	> 7 mmol/l
Normal glycaemia	4 mmol/l < > 7 mmol/l

4 THE APPROACH FOR GLYCEMIAS RECOGNITION AND PREDICTION

For recognising and predicting glycaemias, we utilised the standard machine-learning approach. We constructed the instances which contain a group of extracted features of one or more signals. These

instances are processed using a machine-learning algorithm for classification and the result is evaluated using a 10-fold-cross validation approach.

We evaluated two approaches:

- 1) Single model approach: We first use only the ECG sensor features. In the next steps, we gradually add additional features into the instances and evaluate the recognition and prediction.
- 2) Two-model approach: We first divide the dataset according to the time point the measurement was performed (before or after the meal) and use this information as a context to divide the decision space. We use one model for recognition or prediction of glycaemias before the meal and other for recognition or prediction of glycaemias after the meal.

Both approaches were evaluated in four setups, each setup using different set of attributes for the used signals respectively and being labelled with the current glucose level (hypo-, hyper-, or normal glycaemia). The attribute sets are:

A1: All attributes (absolute and relative values)

A2: Absolute attribute values

A3: Relative attribute values

A4: Top 20 attributes as recommended by the ReliefF algorithm [14]

Figure 3 shows the number and the distribution of glycaemia occurrences in the dataset, where we can see that the most common cases are hyperglycaemias, and that there are only a few cases in total of hypoglycaemias for patients with type I diabetes. No cases of hypoglycaemia were recorded in patients with type II.

Figure 4 shows the number and the distribution of glycaemias when we separate the dataset with respect to the “time of glucose measurement” attribute values “before” or “after meal” for the second approach. We observe that for predicting glycaemia, the only data available for analysis is for diabetes type I before meal.

5 EXPERIMENTS AND RESULTS

We carried out 16 experiments for glycemia prediction and 16 experiments for glycemia recognition. For each set of attributes (A1 to A4 from Section 3.2), we built models with the following approaches:

M1: Model built using attributes from the ECG signal

M2: Model built using attributes from the ECG signal and the respiration rate measurements

M3: Model using both the ECG signal, respiration rate measurements, and the “time of glucose measurement” attribute

M4: Two models are built, each to be used according to the “time of glucose measurement” attribute. One model is built for “before meal” and the second for the “after meal” classification. Both models are built with the same signal data as M2.

Each set of attributes was tested using ten machine-learning algorithms, as implemented in the Weka machine-learning suite [9] using the default algorithm parameters: Naïve Bayes, Logistic Regression, SVM, IB3, AdaBoostM1 with RepTree, Bagging with

RepTree, JRip, J48, Random Forest, and ZeroR as the basic algorithm that always returns the dominant class. Each experiment was evaluated and tested using the 10-fold cross validation.

The results of the algorithm testing are presented in Table 2 for

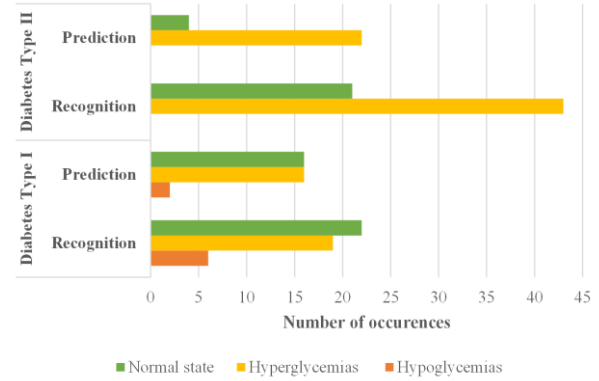


Figure 3. Distribution and number of glycaemias according to diabetes type and classification task.

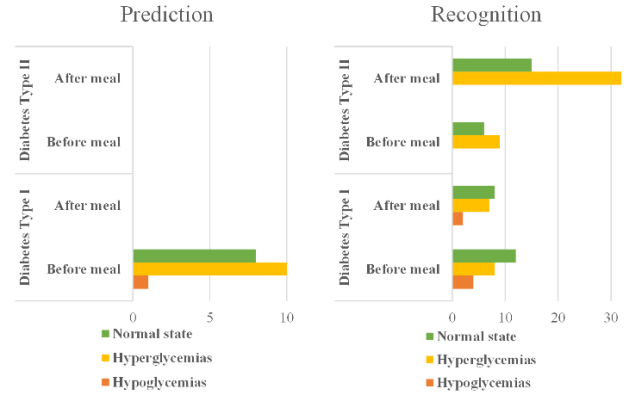


Figure 4. Distribution and number of glycaemias according to classification task, diabetes type and time point of glucose measurement.

the diabetes type I and in Table 3 for diabetes type II.

When predicting glycaemias, the set of attributes A4 always returned best results for diabetes type I patients while the A2 set was best for diabetes type II patients. The highest accuracy for type I diabetes patients, 84 %, was obtained using logistic regression and by separating the dataset based on the glucose measurement time with approach M4. We were unable to evaluate the same approach on type II patients due to the lack of data. We suspect that diabetes type II patients mostly measured their glucose levels when feeling bad, since the measurements were not done before or after meals but at various times throughout the day. The best result for type II was obtained with the IB3 algorithm, with 88 % accuracy.

For glycaemia recognition, the best results were obtained with the M4 approach for both types of diabetes. For type I, the best results were obtained using the A4 set of attributes and the SVM algorithm for the model before meal and logistic regression for after meal. This approach resulted in 78 % accuracy. For type II, the best results were for the A4 set and SVM before meal and A2 and Bagging algorithm after meal. This approach resulted in 76 % accuracy.

Table 2. Recognition and prediction accuracies for the best combination of attributes (A) and approaches (M) for diabetes type I.

	ZeroR (%)	Acc (%)	A	M
Glycaemia recognition	49	78	A4	M4
Glycaemia prediction	53	84	A4	M4

Table 3. Recognition and prediction accuracies for the best combination of attributes (A) and approaches (M) for diabetes type II.

	ZeroR (%)	Acc (%)	A	M
Glycaemia recognition	66	76	*	M4
Glycaemia prediction	85	88	A2	M3

* A4 before meal, A1 after meal

We achieve reasonable accuracies both for recognition and prediction in this preliminary analysis which shows that our approach is promising. However, we should note that the data for predicting the glycaemia in diabetes type II patients was extremely unbalanced, containing 85% of cases of hyperglycaemia and not a single hypoglycaemia, with other measurements being normal state. The results of glycaemia prediction in type II are therefore not representative.

6 CONCLUSION

We present a machine-learning based approach to predict and recognize anomalous blood glucose levels (hypo- and hyperglycaemia) for patients with type I and II diabetes. A general machine-learning approach was used to build classification models, based on attributes obtained from the ECG signals and respiration rate measurements.

Experiments were carried out on 30 patients with type I diabetes and 22 patients with type II. We figured out that the best approach in both recognising and predicting glycaemias is to construct two models, one for before and the other for after the meal.

With our approach, we achieved 84 % accuracy for prediction of glycaemias for patients with type I diabetes. Due to the lack of data, we were not able to use the same approach with type II patients, as they were monitoring their glucose level more sparsely and mostly at time when they felt bad. The same two-model approach returned the best results for recognition of glycaemias, we achieved 78 % in case of diabetes type I and 75 % in case of diabetes type II patients.

The results seem somewhat surprising since one would expect that recognizing glycaemias is easier than predicting them. We plan to investigate this further to better understand it.

In future work, we plan to pre-process the raw data using other types of filtering approaches which will enable us to keep more clean data around glucose measurement time points. We will add additional features such as recognised activities and estimated energy expenditure during the day and other collected data during the pilot study, such as the blood pressure and weight for a more personalised approach. We believe that knowledge about the activities of the patients and the intensity of activity will significantly contribute to more accurate recognitions and predictions of the glycaemias.

Nevertheless, we will also evaluate whether the presented method and future work method is appropriate for practical use,

namely, to actively advise patients to check their glucose levels using their standard (invasive) equipment before symptoms occur.

ACKNOWLEDGEMENTS

The study was partially financed by the EU project COMMODITY12 (www.commodity12.eu).

REFERENCES

- [1] International Diabetes Federation, <http://www.diabetesatlas.org/>
- [2] B. Cvetković, V. Janko, and M. Luštrek, M. Demo abstract: Activity recognition and human energy expenditure estimation with a smartphone. *In PerCom 2015* (2015).
- [3] B. Cvetković, R. Milić, and M. Luštrek, 2016. Estimating Energy Expenditure with Multiple Models using Different Wearable Sensors. *IEEE Journal of Biomedical and Health Informatics*, accepted for publication in July 2016. DOI= 10.1109/JBHI.2015.2458779.
- [4] Commodity12 Project, <http://commodity12.eu/>
- [5] Zephyr BioHarness, <http://www.zephyranywhere.com/>
- [6] GlucoTel, <http://bodytel.com/portfolios/glucoTel/?lang=en>
- [7] M. Hanefeld, E. Duetting, P. Bramlage, Cardiac Implications of Hypoglycaemia in Patients with Diabetes – a Systematic Review. *Cardiovascular Diabetology*, **135**, 12 (2013).
- [8] M.B. Frier, G. Schernthaner, R. Simon, R.S. Heller, Hypoglycemia and Cardiovascular Risks. *Diabetes Care*, **34**, 132-137 (2011).
- [9] J.K. Snell-Bergeon, R.P. Wadwa, Hypoglycemia, Diabetes, and Cardiovascular Disease. *Diabetes Technology & Therapeutics*, **14**, 51–58 (2012).
- [10] P.J. Singh, G.M. Larson, J.C. O'Donnell, F.P. Wilson, H. Tsuji, M.D. Lloyd-Jones, D. Levy, Association of hyperglycemia with reduced heart rate variability (The Framingham Heart Study). *The American Journal of Cardiology*, **86**, 3, 309-312 (2000).
- [11] L.L. Nguyen, S. Su, H.T. Nguyen, Identification of Hypoglycemia and Hyperglycemia in Type 1 Diabetic patients using ECG parameters, in *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society EMBC*, IEEE, 2716-2719.
- [12] K. Plis, R. Bunescu, C. Marling, J. Shubrook, F. Schwartz, A Machine Learning Approach to Predicting Blood Glucose Levels for Diabetes Management. *Workshops at the Twenty-Eighth AAAI Conference on Artificial Intelligence*, 2014.
- [13] E.B. Mazomenos et al., "A Low-Complexity ECG Feature Extraction Algorithm for Mobile Healthcare Applications," in *IEEE Journal of Biomedical and Health Informatics*, **17**, 2, 459-469 (2013).
- [14] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, H.I. Witten, The WEKA Data Mining Software: An Update. *SIGKDD Explorations*, **11**, 1, 2009.

Developing a Motivational System to Manage Physical Activity for Type 2 Diabetes

Yousef Alfaifi, Floriana Grasso and Valentina Tamma¹

Abstract. Type 2 diabetes (T2D), a chronic disease, can be effectively managed with the combination of diabetic medications and a healthy lifestyle. Regular physical activity is an example of a healthy lifestyle that helps to manage T2D and prevents complications. However, barriers to physical activity prevent and hinder diabetic patients from living a healthy lifestyle. Patients' health condition and personal obstacle are common barriers to physical activity. This paper describes preliminary work towards the development of a framework to motivate patients with T2D to engage in regular physical activity. Basic information, current health conditions, and the behaviour of diabetic patients will also be included in the framework for the identification of specific barriers. Insights from persuasive technology will be incorporated into the framework to motivate the patient to healthy lifestyle modification. The framework is based on a model understanding of behaviour and behaviour change of patients.

1 Introduction and Motivation

Diabetes is a complex and chronic disease requiring expensive, psychological treatment, continuous medical care and self-management by the patient [3]. The recent statistics indicate a dramatic increase in the number of diabetic people around the world, reaching 422 million in 2014 compared with only 108 million in 1980 [17]. This number is expected to increase to 552 million by 2030 [21] and 592 million by 2035 [8]. Annually, diabetes is estimated to cost around 10% of the total health budget, and this percent is projected to reach to 17% by 2035 [11]. Diabetes and its complications cause more than two million deaths each year [17]. *Type 2 Diabetes (T2D)* is the most common type of diabetes; approximately 90-95% of all diabetes cases worldwide are T2D [3]. Other types of diabetes include type 1 diabetes, gestational diabetes mellitus and monogenic diabetes syndromes [3]. T2D, also known as "non-insulin-dependent diabetes" occurs when the body cannot use its' insulin effectively [3, 21]. Diabetic medications, either multiple-dose insulin injections or low-dose tablets, and a healthy lifestyle can help manage T2D [3]. Public health professionals have begun focusing increasingly on lifestyle changes to improve the management of T2D and diabetics' overall health [3, 17–19]. A healthy lifestyle can include regular physical activity, nutrition planning, smoking cessation etc [3]. The World Health Organisation (WHO) defines *physical activity* as "any bodily movement produced by

skeletal muscles that requires energy expenditure" [17]. Conversely, unhealthy lifestyles lead to poor health management and increase the risk of developing T2D [3, 19].

Although regular physical activity supports a patient's self-management of diabetes [19], there are barriers and obstacles that prevent patient from achieving the maintain physical activity [20]. These *barriers* can be defined, in general, as obstacles that prevent diabetic patients from living a healthy lifestyle, either partially or totally. Physical activity barriers are usually environmental, personal or medical constraints [14, 20]. Most of these barriers are shared with the non-diabetic population, typically linked to lack of motivation [3, 12]. In addition, there are specific psychological [6] and health barriers for patients with T2D such as an absence of stimulus and hypoglycaemia, respectively [14, 19].

The most recent report from WHO and American Diabetes Association (ADA) suggests that advanced computer technology can support and improve the self-management of diabetes [3, 17]. The technology can improve individual's lifestyles and lead to behaviour changes that support the better management of T2D and prevent or delay T2D development [3]. Moreover, technology can also motivate a patient with regard to better lifestyle modification [18]. This paper presents a preliminary framework to assist patients with T2D to manage the physical activity barriers and persuade to lifestyle modification. Computer technologies that advise or persuade a patient regarding lifestyle modifications are based on a model of patients' behaviour and behaviour change in achieving regular physical activity.

The rest of paper is organised as follows: Section 2, we look at the problem statement. In section 3 debates the literature review. Section 4 discusses the methodology of the framework. Section 5 presents the evaluation of the system. Section 6 gives points of the expected challenges. Finally, a brief conclusion and discussion about future work are given in Section 7.

2 Problem Statement

Healthy lifestyle choices, such as regular physical activity, offer a healthy and economical way to monitor the T2D. Barriers to physical activity are the main problem that obstruct typical physical activity, and therefore a satisfactory lifestyle. Today, computer technology plays a vital role in enabling a patient to overcome complex problems, provide proper advice, and influence a patient to realise positive behaviour modifications. In this proposed framework, we will mainly investigate the opportunity of computer technology intervention (rule-based system and persuasive technologies) in managing

¹ Computer Science department, Liverpool University, UK, email: Y.Alfaifi, Floriana, V.Tamma@liverpool.ac.uk

ing physical activity and motivation to lifestyle changes for T2D patients. The model patient's behaviour and behaviour changes will be taken into consideration to ensure a convincing investigation and ensure we overcome the main problem (barriers to physical activity). Identifying the barriers to physical activity based on features or signs are presented as central roles in addressing the issue. Therefore, judging and assessing the ability of a rule-based system, particularly *if-then* rules, to identify accurate barriers to physical activity will be reviewed initially. An evaluation and estimation of the strength of persuasive technology to persuade and influence patient modification will be examined as well. As we will mention in Section 4, depending on physical activity barriers, the proposed system provides suitable advice at the end. Accordingly, the ability of the *If-then* formula to suggest correct advice will be measured and tested, too. Evaluating and studying these combined cases can guide this research to explore the capability of computer technology to manage the barriers to physical activity for T2D patients.

3 Literature Review

Computer technology, including rule-based systems, has been successfully utilised within the field of professional healthcare to develop the health services that are provided to patients [10]. Rule-based systems, or expert systems, which are a branch of artificial intelligence, are employed in the field of medicine to support self-management, advising, decision-making and support. A *rule-based system* results in the transfer of a human experience or human knowledge into an automated system in order to solve complex problems. A rule-based system is based on a set of facts and rules that uses *if-then* statements to make a decision-making [7].

In the diabetes field, various types of expert system have been improved to support diabetes patients in managing the disease. Some of the diabetes studies are discussed below.

The authors in [16] demonstrate a diabetes management model to enable patients with T2D to alter their lifestyles. The goal of this model is to monitor and interpret a patient's daily lifestyle changes in the form of decision support to achieve patient's health goals. Seven inputs are necessary to insert into the model: age, gender, weight, height, blood glucose (BG), exercise, and diet. The system returns three outputs: glycatedhaemoglobin (A1c), exercise and diet level assessments.

The study in [1] designed and implemented a rule-based expert system to manage one type of lifestyle, which is a healthy diet for patients with T2D. The system can provide the patient with a plan for a satisfactory amount of daily calories as well as a list of proposed foods.

The study in [9] developed an approach to support patients in the self-management of diabetes. The approach is based on an expert system to represent the knowledge. The system includes four sections: body weight and the assessment of daily nutritional requirements, hypoglycaemia symptoms, self-monitoring of blood glucose levels and diabetes-related disease. The system produces appropriate recommendations to manage diabetes by using the patient's input, including blood glucose levels and other related signs.

The authors in [13] constructed incorporating diet in the self-management of diabetes. They use a case-based approach to advice patients about healthy nutrition. The case-based approach acquires its knowledge from clients and patients.

The rule-based system included in this approach uses a health record management system. The system filters the food products while considering the patient's or client's health record. The system can ultimately produce a suggested diet plan.

The authors in [10] established a system to advise women with gestational diabetes regarding the adjustment of daily multiple-dose insulin and dietary habits. The system provides a consultation according to patient inputs, which include blood glucose level, time and nutrition modification. The system is evaluated in real scenarios and has proven to reduce the frequency of doctor visits.

The above studies effort in diabetes management, but are lacking when it comes to addressing patient behaviours that can significantly impact the management of their disease. In order to take these behaviours into consideration, the behaviour of user need to be understood [5, 15]. Once user behaviour can be appreciated, system developers are able to create a motivational system that has the ability to change user behaviour [5], rather than just provide a simple consultation. Conversely, a system that is designed without user behaviour in mind will yield a highly limited solution [5]. A patient or user may know, obviously, that eating healthier food leads to a healthier lifestyle, and vice versa, but the results are apparent in the future, not immediately. Imagine using a short video to show the direct cause-and-effect relationship between nutritional eating and a healthy or unhealthy lifestyle, and how this would affect the behaviour of the patient. This simulation lets the user explore and experiment with a real healthy or unhealthy consequence [4]. The simulation, which is based on motivation factors, simulates particular pleasure and pain elements and pushes a user to change the behaviour [4, 5]. Fogg describes this technique as *persuasive technology*, which is defined as "learning to automate behaviour change" [5].

Psychological research studies have shown that opportunities for learning behaviour changing techniques, such as motivation and goals, influence a person's behaviour modification [15]. Motivating a diabetic patient to change the lifestyle, like quitting smoking, is more efficient than just treatment alone [3]. According to the national standards for Diabetes Self-Management Education, diabetic patients must understand that a healthy lifestyle begins with high-quality self-management to improve overall health and prevent complications of T2D [3]. But how do we encourage, promote, and convince them to act on their beliefs? Consequently, a substantial problem is finding ways we can influence and persuade diabetic patients to follow a healthy lifestyle as directed through the medical advice. Computer technology can play a motivational role in persuading patients to change their behaviour, despite a low health status [4, 5, 18]. In order to effectively apply technology to influence a patient's behaviour change, the patient's behaviour must be taken into consideration [18].

Today, it has become possible to insert persuasive technology into the system design to persuade users to change the behaviour [4, 5]. The Fogg Behaviour Model (FBM) combines the psychological and technological sides in order to push a user towards behaviour modification. FBM is a suitable model to apply, in part, to this framework. FBM paves the way for the movement and application of the psychological theory to computer technology to influence user behaviour modification. FBM is a general model which can be used in the healthcare field to modify patient's behaviour. FBM as-

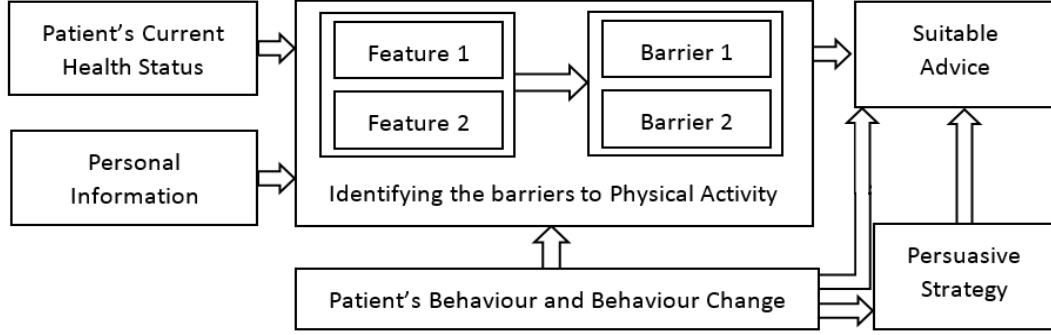


Figure 1. Preliminary Framework to Manage the Physical Activity for T2D

sports that there are three combined factors (sufficient motivation, ability and trigger), which have to come together at the same time for a target behaviour to happen; otherwise, the behaviour will not occur. These factors have provided a platform for designers and researchers to understand users' behaviour and performance [5].

4 Methodology

4.1 Preliminary Framework for Managing Physical Activity

This research presents a preliminary framework for managing physical activity in individuals with T2D (Fig. 1). The framework is based primarily on a model of a patients' behaviour and behaviour change, to capture the actual barriers, to provide a final exhortation, and to design the persuasive strategy. The personal information phase has enabled the system to obtain the necessary information from a patient such as age, gender, city, job (part-time or full-time) and other information. This phase assists in identifying basic features of barriers, such as lack of time, in the early stages. e.g, according to a patient's daily diary, the time constraint barriers can appear, in part, in this phase. The phases of patient's behaviour and behaviour change are based on the psychological theory and Fogg's model to complete this phase successfully. Emotions, social influences, motivations and goals (and other aspects) should be determined in this phase, as well as a patient's beliefs about their capabilities and consequences. These key determinants identify not only the features of psychological barriers, but also help to understand a patient's behaviour and behaviour change [15]. The phase of a patient's current health status must be examined to identify any barriers related to their current health condition. Based on the patient's input, e.g, blood glucose level, and blood pressure, the system can identify whether hypoglycaemia or high blood pressure, respectively, are barriers to physical activity or not. Identification of a patient's physical ability, in this stage, also helps to produce a suitable intensity, type and duration of physical activity by the end of the advising. The phase of the persuasive strategy is based on the FBM, particularly on persuasive technology for behaviour modification, as well as the behaviour and behaviour change of a model patient. The phase of identifying the physical activity barriers is responsible for recognising the actual barriers to physical activity based on the features of the barriers, from either the other phases or its own features (Subsection 4.2). Finally, a correct motivational advice phase can produce a stimulating advice depending on all of the above phases.

The preliminary model presents how these related phases (the personal information, the patient's current health status, the identification of the barriers to physical activity, and the persuasive strategy) can produce suitable advice for diabetic patients depending on the behaviour and behaviour change of a model patient (Fig. 1).

4.2 Proposed Method to Identify the Actual Barriers to Physical Activity

In the preliminary framework on Fig. 1) particularly, in the phase of identifying the barriers to physical activity, we suggest dealing with each barrier as an independent problem. Consequently, each barrier will be identified according to its own features or signs. A *feature* or a *sign* is an attribute or aspect of a barrier to physical activity. One or more features can identify the actual barriers. Ignoring these features, in decision-making about barriers, may lead to incomplete advice, or worse, incorrect advice at the end. Identifying the barriers with certainty guides advisors to successful and suitable advising at the end. i.e, bad weather is presented as a barrier for diabetic patients, as well as the general population [12]. Classifying either the weather condition as a barrier or not is dependent upon certain related states (signs) of the atmosphere and phenomena such as heat, cold, storms, and rain. The evaluation or assessment of each factor acts as a guide for accurate decision-making and as a recommendation with more details, whether on the barriers side or on the advising side. In contrast, ignoring one or more these signs or features, even though they are forecasted in some cases, leads to inexact identification of barriers, and, consequently, inaccurate advice. The below *if-then* formula clarifies how to identify if the weather is a barrier (or not), based on a few weather signs.

*IF it is winter, OR the temperature is < 0 degrees
OR the weather is stormy
THEN the weather is a barrier because it is cold, we advise
you to do indoor physical activity*

On the side of health barriers, hypoglycaemia, or low blood glucose, is classified as an obstacle to maintaining physical activity for the patient with T2D [14, 19]. Symptoms such as hunger and nausea, blurred (impaired) vision, and a headache will be present with hypoglycaemia [2]. The blood glucose level, an indicator of the patient's current health status, is also examined in hypoglycaemia [19]. The following *if-then* rules explain how to determine whether hypoglycaemia is a barrier to physical activity or not, based on a few symptoms (features) [2, 19]. It also displays a caution when performance of physical activity is inadvisable due to potential side effects.

IF blood glucose levels < 100 mg/dL (5.6 mmol/L)
 OR feeling hungry, OR blurred/impaired vision
 THEN stop physical activity, recheck your blood, glucose
 after 15 minutes, and eat a small snack

5 Evaluation

The evaluation of the proposed system includes two main stages. The first stage involves evaluating and estimating the ability of the proposed rules to identify specific problems (i.e., barriers to physical activity) by using a *forward-chaining mechanism*. Forward chaining can match a patient's input (barrier's feature) to decide which rules are fired and then provide advice derived from the data. Forward chaining show the capability of the rules to identify either the weather or hypoglycaemia, respectively, as barriers based on certain features and symptoms (Subsection 4.2). The second stage will be the evaluation of the entire system. Patients with T2D, health care providers, and specialists would contribute to system evaluation. Feedback will be taken into consideration to improve the proposed system.

6 Expected Challenges

Academic researchers can expect to face challenges in any area of investigation. Anticipating challenges and seeking suitable solutions in the early stages of research serves to help the researcher manage difficulties more efficiently. The anticipated challenges of this study include:

- Identify the specified barriers based on several features of barriers (psychological, medical or personal), and then produce suitable advice.
- Modeling of patient's behaviour and behaviour change in different age groups, and designing the persuasive strategy, e.g., persuasive technology, with these differences in mind.

7 Conclusion and Future Work

Helping patients with T2D perform regular physical activity to result in lifestyle modifications is a challenge faced by health organisations and researchers. At the individual level, a patient regularly partaking in physical activity contributes to the maintenance of a healthy lifestyle and in assistance with T2D management, however barriers often prevent meaningful physical activity. The framework described in this paper proposes a system by which to manage barriers to physical activity, improving lifestyle changes, and supporting T2D management. Both rule-based system (*if-then* rules) and persuasive technologies integrate with this framework, which works to identify physical activity barriers and providing correct advice at the end. Developing, testing and additional evaluation of the preliminary framework will be conducted in future work.

Diabetes is only one of many chronic conditions impacting peoples' lives. The preliminary proposed framework can be applied to different chronic diseases, including type one diabetes, obesity and high blood pressure. The method of identifying physical activity barriers according to features can also be applied to other chronic diseases.

REFERENCES

- [1] Ibrahim M Ahmed and Abeer M Mahmoud. Development of an expert system for diabetic type-2 diet. *Development*, 107(1), 2014.
- [2] American Diabetes Association. Hypoglycemia (low blood glucose), 2016. <http://www.diabetes.org/living-with-diabetes/treatment-and-care/blood-glucose-control/hypoglycemia-low-blood.html>.
- [3] American Diabetes Association et al. Standards of medical care in diabetes-2016. *Diabetes care*, 39(Supplement 1):S1–S112, 2016.
- [4] Brian J Fogg. Persuasive technology: using computers to change what we think and do. *Ubiquity*, 2002(December):5, 2002.
- [5] Brian J Fogg. A behavior model for persuasive design. In *Proceedings of the 4th international Conference on Persuasive Technology*, page 40. ACM, 2009.
- [6] Russell E Glasgow, Deborah J Toobert, and Cynthia D Gillette. Psychosocial barriers to diabetes self-management and quality of life. *Diabetes spectrum*, 14(1):33–41, 2001.
- [7] Crina Grosan and Ajith Abraham. Rule-based expert systems. In *Intelligent Systems*, pages 149–185. Springer, 2011.
- [8] L Guariguata, DR Whiting, I Hambleton, J Beagley, U Linenkamp, and JE Shaw. Global estimates of diabetes prevalence for 2013 and projections for 2035. *Diabetes research and clinical practice*, 103(2):137–149, 2014.
- [9] Baran Hashemi and Hossein Javidnia. An approach for recommendations in self management of diabetes based on expert system. *International Journal of Computer Applications*, 53(14), 2012.
- [10] M Elena Hernando, Enrique J Gómez, R Corcoy, and Francisco del Pozo. Evaluation of diabnet, a decision support system for therapy planning in gestational diabetes. *Computer methods and programs in biomedicine*, 62(3):235–248, 2000.
- [11] N Hex, C Bartlett, D Wright, M Taylor, and D Varley. Estimating the current and future costs of type 1 and type 2 diabetes in the uk, including direct health costs and indirect societal and productivity costs. *Diabetic Medicine*, 29(7):855–862, 2012.
- [12] Eveliina E Korkiakangas, Maija A Alahuhta, Päivi M Hussian, Sirkka Keinänen-Kiukaanniemi, Anja M Taanila, and Jaana H Laitinen. Motivators and barriers to exercise among adults with a high risk of type 2 diabetes—a qualitative study. *Scandinavian journal of caring sciences*, 25(1):62–69, 2011.
- [13] Gergely Kovasznai. Developing an expert system for diet recommendation. In *Applied Computational Intelligence and Informatics (SACI), 2011 6th IEEE International Symposium on*, pages 505–509. IEEE, 2011.
- [14] Julia Lawton, N Ahmad, L Hanna, M Douglas, and N Hallowell. 'i can't do any serious exercise': barriers to physical activity amongst people of pakistani and indian origin with type 2 diabetes. *Health Education Research*, 21(1):43–54, 2006.
- [15] Susan Michie, Marie Johnston, Jill Francis, Wendy Hardeman, and Martin Eccles. From theory to intervention: mapping theoretically derived behavioural determinants to behaviour change techniques. *Applied psychology*, 57(4), 2008.
- [16] Nonso Nnamoko, Farath Arshad, David England, Jiten Vora, and James Norman. Fuzzy inference model for type 2 diabetes management: a tool for regimen alterations. *Journal of Computer Sciences and Applications*, 3(3A):40–45, 2015.
- [17] World Health Organization. Global report on diabetes. *Diabetes research and clinical practice*.
- [18] Benjamin A Rosser, Kevin E Vowles, Edmund Keogh, Christopher Eccleston, and Gail A Mountain. Technologically-assisted behaviour change: a systematic review of studies of novel technologies for the management of chronic illness. *Journal of Telemedicine and Telecare*, 2009.
- [19] Ronald J Sigal, Glen P Kenny, David H Wasserman, and Carmen Castaneda-Sceppa. Physical activity/exercise and type 2 diabetes. *Diabetes care*, 27(10):2518–2539, 2004.
- [20] N Thomas, E Alder, and GP Leese. Barriers to physical activity in patients with diabetes. *Postgraduate Medical Journal*, 80(943):287–291, 2004.
- [21] David R Whiting, Leonor Guariguata, Clara Weil, and Jonathan Shaw. Idf diabetes atlas: global estimates of the prevalence of diabetes for 2011 and 2030. *Diabetes research and clinical practice*, 94(3):311–321, 2011.

Increasing transparency of recommender systems for type 1 diabetes patients

John Paul Vargheese¹, Rachel Harrison¹, Mireya Munoz Balbontin¹, Arantza Aldea¹, Daniel Brown¹

Abstract. Self-management of type 1 diabetes is a challenging and complex task due to the constant need for self monitoring and the diverse range of factors to consider in order to effectively regulate blood glucose levels. Recommender systems have been demonstrated to be effective for supporting patient self-management of type 1 diabetes by providing recommendations for insulin doses. Recent studies have expanded on this approach by incorporating case based reasoning within existing recommender systems for type 1 diabetes, to provide a more flexible and personalised approach to making recommendations. However, recommendations made by such systems may be ignored, even when users consider the system's performance to be good. To address this, we propose a complimentary approach to increase the transparency of such systems through the provision of explanatory summaries that expose the reasoning process for making the recommendation. Greater transparency may increase recommendation acceptance rates and improve users' trust and acceptability of these systems.

1 Introduction and motivation

Type 1 diabetes is an autoimmune disease in which the pancreas is unable to produce insulin which prevents regulation of blood glucose levels (BGL). Regulating optimal BGL is essential in order to avoid severe long-term health problems caused by hyperglycaemia (high blood sugar levels) and hypoglycaemia (low blood sugar levels). Current treatment involves administering insulin which can be delivered either through subcutaneous injections or through an insulin pump. Self-management of type 1 diabetes typically involves the monitoring of BGL using a blood glucose meter and estimating the required amount of insulin to regulate BGL. However, this usually results in a less than optimal regulation of BGL [8]. This combined with the wide range of subjective and individual physiological factors that may affect BGL such as stress, illness, exercise and other activities of daily living and lifestyle, make self-management and treatment recommendations a complex and challenging task [21]. Furthermore, maintaining an optimal self-management regimen can be difficult to achieve due to the need for persistent monitoring of BGL, calculating and administering required insulin doses and following recommendations for increasing exercise and adopting a new healthier lifestyle [14]. Despite these challenges, effective self-management has been demonstrated to be effective for avoiding long term health risks associated with type 1 diabetes [9]. Recommender systems such as insulin bolus calculators support patient

self-management by recommending insulin doses [18, 10] and have been demonstrated to be effective across a range of studies [12, 1]. However, recommendations made by such systems may often require amendments due to the wide variety of factors that may impact upon BGL [15]. To address this, recent studies have demonstrated the benefits and effectiveness of enhancing insulin bolus calculators by incorporating case based reasoning (CBR) which offers a means of providing more flexible and personalised recommendations utilising a knowledge base of previous experiences [15, 2, 8, 3]. However, as in the case of other recommender, knowledge and expert based systems, users sometimes ignore and reject recommendations due to a lack of transparency [7, 13] even in cases where users consider the system's performance to be good [16]. Increasing transparency of such recommender systems by providing an explanatory summary that exposes the reasoning process for a proposed recommendation may increase acceptance rates and improve users' trust and acceptability. Previous work has demonstrated how transparency of recommender systems can increase user trust and acceptability of such systems [4, 5].

2 Research challenges and proposed studies

Hypothesis: To realise the potential benefits of increasing the transparency of recommender systems for type 1 diabetes, a number of research challenges must be addressed. Our hypothesis (*H*) is: Increasing transparency by providing explanations for recommendations will increase acceptance rates, users' trust and acceptability.

Study design: To assess this, it is necessary to consider what metrics to apply to measure these outcomes. For example, consider a preliminary controlled evaluation consisting of two groups of patients, where both groups are provided with sample data from which a recommendation is proposed. Group *A* are provided with a recommendation and no explanation and group *B* are provided with a recommendation with an explanation.

Metrics: Participants are provided with a questionnaire to indicate whether they would accept a recommendation, how much they trust it, whether it reduces the effort for deciding whether to accept or reject a recommendation (using Likert scales) and whether they would consider future recommendations proposed by the system. These outcome measures provide an initial assessment of *H*, however, we propose an iterative series of evaluations varying the strategy for presenting a recommendation. These strategies [20] include *Top recommendation*: Providing a simple explanation for a proposed dose, for example reporting BGL only as part of the recommendation. *Predicted recommendation*: Providing an indication of the user's predicted BGL for accepting a recommendation and for rejecting a recommendation. *Structured overview*: Providing an overview all factors that have been considered by the underlying CBR for example, BGL, glycemic index and physical activity. Further considerations

¹ Oxford Brookes University, UK,
jpvargheese@acm.org
rachel.harrison@brookes.ac.uk
mireya.munoz.balbontin-2016@brookes.ac.uk
aaldea@brookes.ac.uk
dbrown@brookes.ac.uk

for how to present explanations for recommendations and user system interaction, include investigating how varying visualisation options may help to improve the effectiveness and comprehensibility of an explanation. Various studies have demonstrated how visualisations of medical data does not always enhance decision making [6] and is typically most beneficial for expert users but less beneficial for those with varying ranges of expertise [19]. Similarly, visualisation alone has been demonstrated to be less effective for supporting decision making compared to expert authored textual summaries [11].

3 Automating explanations for recommendations

We propose investigating the potential for using natural language generation (NLG) for producing explanations for recommendations. NLG systems analyse data to produce human readable text using a four stage process as shown in Figure 1. The proposed platform in Figure 1 incorporates a standard NLG architecture proposed in [17], that is capable of receiving data and or knowledge as inputs to the system.

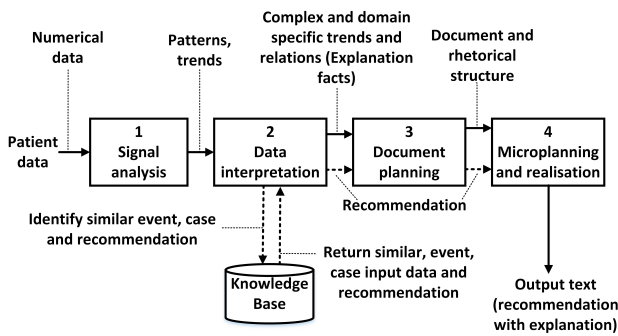


Figure 1. Potential platform for automated explanations incorporating an NLG architecture adapted from [17]

4 Discussion

Recommender systems such as those mentioned in this paper have the potential to significantly reduce the risks associated with type 1 diabetes by supporting patient self-management. Users' trust and acceptance are crucial to ensure widespread use and adoption of such systems. In this paper, we propose investigating a complimentary approach to existing recommender systems for type 1 diabetes patients, by increasing the transparency of recommendations by providing an explanatory summary of a proposed recommendation. We believe this research has the potential to increase acceptance rates, users' trust and acceptability of such systems and may provide insights for developing new models of trust utilising provenance which could potentially enhance the reasoning process for making recommendations.

REFERENCES

- [1] Katharine Barnard, Christopher Parkin, Amanda Young, and Mansoor Ashraf, 'Use of an automated bolus calculator reduces fear of hypoglycemia and improves confidence in dosage accuracy in patients with type 1 diabetes mellitus treated with multiple daily insulin injections', *Journal of diabetes science and technology*, **6**(1), 144–149, (2012).
- [2] D. Brown, *Temporal Case-based Reasoning for Insulin Decision Support*, Ph.D. dissertation, Department of Computing and Communication Technologies, Oxford Brookes University, 2015.
- [3] D. Brown, I. Bayley, R. Harrison, and C. Martin, 'Developing a mobile case-based reasoning application to assist type 1 diabetes management', in *2013 IEEE 15th International Conference on e-Health Networking, Applications & Services (Healthcom)*, Lisbon, Portugal, (2013). IEEE.

- [4] Li Chen and Pearl Pu, 'Trust building in recommender agents', in *Proceedings of the Workshop on Web Personalization, Recommender Systems and Intelligent User Interfaces at the 2nd International Conference on E-Business and Telecommunication Networks*, pp. 135–145. Citeseer, (2005).
- [5] Alexander Felfernig and Bartosz Gula, 'An empirical study on consumer behavior in the interaction with knowledge-based recommender applications', in *E-Commerce Technology, 2006. The 8th IEEE International Conference on and Enterprise Computing, E-Commerce, and E-Services, The 3rd IEEE International Conference on*, pp. 37–37. IEEE, (2006).
- [6] Yvonne Freer, Lindsey Ferguson, Gary Ewing, Jim Hunter, Robert Logie, Sue Rudkin, and Neil McIntosh, 'Mismatched concepts in a neonatal intensive care unit (nicu): further issues for computer decision support?', *Journal of clinical monitoring and computing*, **17**(7-8), 441–447, (2002).
- [7] M Sinan Gönül, Dilek Önköl, and Michael Lawrence, 'The effects of structural characteristics of explanations on use of a dss', *Decision Support Systems*, **42**(3), 1481–1493, (2006).
- [8] Pau Herrero, Peter Pesl, Jorge Bondia, Monika Reddy, Nick Oliver, Pantelis Georgiou, and Christofer Toumazou, 'Method for automatic adjustment of an insulin bolus calculator: In silico robustness evaluation under intra-day variability', *Computer methods and programs in biomedicine*, **119**(1), 1–8, (2015).
- [9] Alan M Jacobson, Barbara H Braffett, Patricia A Cleary, Rose A Gubitosi-Klug, Mary E Larkin, DCCT/EDIC Research Group, et al., 'The long-term effects of type 1 diabetes treatment and complications on health-related quality of life a 23-year follow-up of the diabetes control and complications/epidemiology of diabetes interventions and complications cohort', *Diabetes care*, **36**(10), 3131–3138, (2013).
- [10] David C Klonoff, 'The current status of bolus calculator decision-support software', *J Diabetes Sci Technol*, **6**(5), 990–994, (2012).
- [11] Anna S Law, Yvonne Freer, Jim Hunter, Robert H Logie, Neil McIntosh, and John Quinn, 'A comparison of graphical and textual presentations of time series data to support medical decision making in the neonatal intensive care unit', *Journal of clinical monitoring and computing*, **19**(3), 183–194, (2005).
- [12] G Lepore, AR Dodesini, I Nosari, C Scaranna, A Corsi, and R Trevisan, 'Bolus calculator improves long-term metabolic control and reduces glucose variability in pump-treated patients with type 1 diabetes', *Nutrition, Metabolism and Cardiovascular Diseases*, **22**(8), e15–e16, (2012).
- [13] Saad Mahamood, Ehud Reiter, and Chris Mellish, 'Neonatal intensive care information for parents? an affective approach', in *21st IEEE International Symposium on Computer-Based Medical Systems*, pp. 461–463. IEEE, (2008).
- [14] Cindy Marling, Matthew Wiley, Razvan Bunescu, Jay Shubrook, and Frank Schwartz, 'Emerging applications for intelligent diabetes management', *AI Magazine*, **33**(2), 67, (2012).
- [15] Peter Pesl, Pau Herrero, Monika Reddy, Nick Oliver, Desmond G Johnston, Christofer Toumazou, and Pantelis Georgiou, 'Case-based reasoning for insulin bolus advice evaluation of case parameters in a six-week pilot study', *Journal of diabetes science and technology*, 1932296816629986, (2016).
- [16] Frank Puppe, Martin Atzmueller, Georg Buscher, Matthias Huettig, Hardi Luehrs, and Hans-Peter Buscher, 'Application and evaluation of a medical knowledge system in sonography (sonoconsult)', in *ECAI*, volume 8, pp. 683–687, (2008).
- [17] Ehud Reiter, 'An architecture for data-to-text systems', in *Proceedings of the Eleventh European Workshop on Natural Language Generation*, pp. 97–104. Association for Computational Linguistics, (2007).
- [18] Signe Schmidt and Kirsten Nørsgaard, 'Bolus calculators', *Journal of diabetes science and technology*, **8**(5), 1035–1041, (2014).
- [19] Yuval Shahar, Dina Goren-Bar, David Boaz, and Gil Tahan, 'Distributed, intelligent, interactive visualization and exploration of time-oriented clinical data and their abstractions', *Artificial intelligence in medicine*, **38**(2), 115–135, (2006).
- [20] Nava Tintarev and Judith Masthoff, 'Explaining recommendations: Design and evaluation', in *Recommender Systems Handbook*, 353–382, Springer, (2015).
- [21] Donald Walker, *Similarity determination and case retrieval in an intelligent decision support system for diabetes management*, Ph.D. dissertation, Ohio University, 2007.

Assessment of diabetic complications based on series of records

Eva Armengol¹

Abstract.

We propose an approach to assess the risk of complications for diabetic patients. This assessment is based on previous records of the same patient and also in both records and evolution of similar patients.

Keywords: Diabetes Mellitus, Individual Prognosis, Artificial Intelligence, Case-based Reasoning

1 Introduction

In 1989 held in St. Vincent (Italy) a meeting focused to find ways to improve the health of people with diabetes in Europe. The result of the meeting was the so-called *St. Vincent Declaration* whose basic demands were the use of evidence-based treatment, equity of access and strong partnerships in care for people with diabetes [6, 7].

Diabetes mellitus (DM) is a metabolic disorder in which the human body has not enough insulin to move the glucose into the blood. There are two major types of diabetes: diabetes Type I (or insulin-dependent) usually found in people being less than 40 years old; and diabetes Type II (or noninsulin-dependent) often developed in people over this age. Both forms of diabetes produce the same short-term symptoms (i.e. increase of thirst, and high blood glucose values) and long-term complications. Physicians classify diabetic complications in two groups [5]: 1) Macro-complications: ischemic cardiopathy, low extremities vasculopathy, and stroke; 2) Micro-complications: nephropathy, retinopathy and polyneuropathy

These complications can be delayed or minimized by maintaining the glucose levels in blood close to the ones of a person without diabetes [10]. The prediction of the individual risk to develop long-term complications is based on the analysis of a large quantity of data that have to be continuously evaluated. The therapeutic goals to offer a good life quality to the patient depend on this analysis.

The DIABCARE Q-Net project [12] developed a complete and integrated information technology system to monitor diabetes care, according to the gold standards of the St. Vincent Declaration Action Program. Inside this project, partners developed what they call a *basic information sheet* that contains around 150 items about a diabetic patient. These items are basic patient data, risk factors, and blood analysis results, in addition to other general information such as the ability of the patient to monitor himself, results of eye and limb examinations, etc. Based on this information, we developed DIRAS [11] (Diabetes Individualized Risk Assessment System), an application whose goal is to predict the risk of complications for diabetic patients.

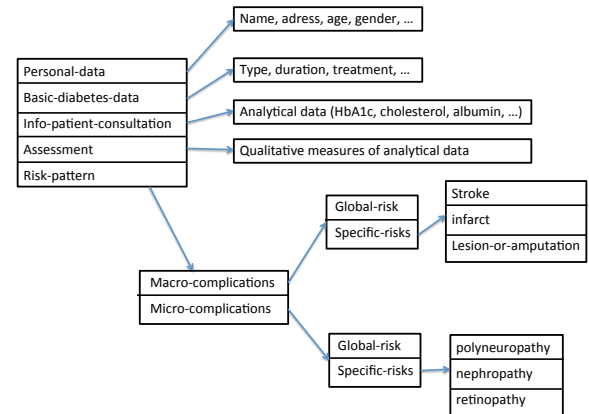


Figure 1. Features describing a diabetic patient in DIRAS.

In the present paper we want to improve DIRAS by taking into account several clinical sessions of each patient to analyse the evolution of that patient when assessing the risk of complications. Most of works on diabetes management are oriented to a global approach, taking into account the characteristics of a given patient and proposing an appropriate insulin dose to maintain blood glucose levels in the normal range (see [3, 4, 14, 15] among others). Our goal is to make a global assessment of the complications, based on the historical record of a patient but also on similar historical records of other patients. In the next section DIRAS is briefly described and then the new approach to be constructed on top of DIRAS is introduced.

2 DIRAS

DIRAS is an application oriented to support the physicians to determine the risk pattern for each diabetic patient according to the clinical data of that patient. The outcome of DIRAS is a *risk pattern*, i.e., a set of assessments concerning diabetic complications. The main contribution of DIRAS was to focus on individual patients instead of populations of patients. For each patient, DIRAS works with five kinds of data (Fig.1): Personal-Data, Basic-Diabetes-Data, Info-Patient-Consultation, Assessment, and Risk-Pattern. Personal-data contains information such as the name, address, birth date, etc. Basic-Diabetes-Data contains basic information of diabetes (such as diabetes type, duration, and whether diabetes is treated with oral drugs

¹ Artificial Intelligence Research Institute (IIIA-CSIC), Bellaterra, Catalonia, Spain, email: eva@iiia.csic.es

```

Function LID ( $p, S_{D_i}, D_i, C$ )
  if stopping-condition( $S_{D_i}$ )
    then return  $class(S_{D_i})$ 
  else  $f_d := \text{Select-attribute}(p, S_{D_i}, C)$ 
        $D_{i+1} := \text{Add-attribute}(f_d, D_i)$ 
        $S_{D_{i+1}} := \text{Discriminatory-set}(D_{i+1}, S_{D_i})$ 
       LID ( $p, S_{D_{i+1}}, D_{i+1}, C$ )
  end-if
end-function

```

Figure 2. The LID algorithm: p is the problem to be solved, D_i is the similitude term, S_{D_i} is the discriminatory set associated with D_i , C is the set of solution classes, $class(S_{D_i})$ is the class $C_i \in C$ to which all elements in S_{D_i} belong.

or insulin). Info-Patient-Consultation has data on relevant measures (e.g. glycated hemoglobin, cholesterol, blood pressure, etc), eye and foot examination, current treatments, etc. Assessment contains qualitative assessments of the data in Info-Patient-Consultation. Risk-Pattern is the assessment of individual long-term risks of a patient. The Risk-Pattern has two parts: 1) the macro-complication risks, and 2) the micro-complication risks. There are two kinds of risk for complications: development risk and progression risk. The development risk has to do with patient's likelihood of developing a new complication in the future. The progression risk is when a patient already has a complication and thus the risk of further deterioration has to be assessed. The goal of DIRAS is to obtain an individual risk pattern for diabetic patients using LID [1] a Case-based Reasoning (CBR) [9] method. DIRAS obtains the risk for each feature in an independent way.

The LID method (*Lazy Induction of Descriptions* algorithm) is shown in Figure 2. The basic idea is to start with a patient description, namely D_i that is the most general one (i.e., an empty description satisfied by all the patients in the case base) and to specialize it by adding features until reaching a description D'_i satisfied by cases that belong to the same solution class. In our diabetes domain, the cases in the case base are complete, in the sense that they have the risk pattern filled with the corresponding assessments. The new problem p has not the Risk-pattern. The features added to specialize a description D_i are added with the value that the feature holds in p . For instance, if the feature selected to add to D_j is *albumin* and p has *albumin* with value *high*, the current D_j should be specialized by adding the feature $\langle \text{albumin}, \text{high} \rangle$. The version of DIRAS introduced in [1] does not take into account the features in Info-Patient-Consultation to compare cases. The solution classes are independent for each feature of the risk pattern, and the labels are *low*, *medium*, *high*, *very high* for each one of them. Therefore, if we are assessing the risk of p for retinopathy, LID will stop if all the cases satisfying the current D_j have the same risk for retinopathy without being aware of the risk of the other complications.

3 Assessing the Risk of Complications based on Similar Historical Records

Diabetic patients are periodically controlled by a physician. During the visits, in addition to the results of clinical analysis, the doctor also inspects eyes, limbs, and asks for the life style of the patient. Such visit gives a picture of the patient's state at that moment. This picture is the one registered in the features Personal-data, Basic-diabetes-

data and Info-patient-consultation show in Fig 1. DIRAS uses domain knowledge to give qualitative values to those in Info-patient-consultation and to fill, in that way, the features in Assessment. Such qualitative description of the patient is the one considered by DIRAS to search for similar patients and to assess the complication risks. Our point is that the risk could be better assessed by taking into account historical records of that patient. For instance, a high value of LDL-cholesterol is more significant when it has been historically high, than when it is the first time that it is high. Our proposal is to qualitatively analyze the records of a given patient to capture the evolution in terms of some linguistic variables such as *low*, *normal* or *high*.

Figure 3 shows an example of a patient's record corresponding to 7 visits. In that record the physician can see that the HDL-cholesterol has been always normal but that the LDL-cholesterol is now high although in the most of previous visits, with only one exception, it has been normal. The patient has high levels of both HbA1c and blood pressure and the creatinine has decreased to normal levels in the last visits (this could mean that the applied therapy is effective). Concerning physical aspects, the patient has no problem in eyes, but there is some abnormality in the left leg.

LDL-cholesterol	N	N	H	N	N	N	H
HDL-cholesterol	N	N	N	N	N	N	N
HbA1c	L	N	H	H	N	H	H
Creatinine	H	H	H	H	N	N	N
Blood pressure	H	H	H	H	H	H	H
Left eye	N	N	N	N	N	N	N
Right eye	N	N	N	N	N	N	N
Left leg	N	N	N	N	N	A	A
Rigth leg	N	N	N	N	N	N	N
	1	2	3	4	5	6	7

Figure 3. Example of a historical record of a patient. It registers 7 visits and the result of the examinations of several features.

This is a very preliminary work, so there are several issues that still have to be fixed. In the next sections we discuss some of them.

3.1 Patient representation

We have available records of patients, with the information we need to apply our approach. We estimate that there are not many records for a given patient since, for instance, patients with Type II diabetes are mostly elder people that commonly have one control at year. In principle, this would not be a shortcoming, and this is the main reason that suggest us to use qualitative data instead of numerical data. Large series of numerical data give us curves that could be analyzed using standard methods such as [8, 13] among others. Maybe this could be appropriate for patients with diabetes type I, so we need to analyze this point in more depth once we have data available. We think that for short historical series of records, it could be easiest to use a qualitative assessment of measures. Therefore, in a preliminary study we will use a qualitative representation of patients as in DIRAS. This means that the measures in Info-patient-consultation will

be discretized using domain knowledge and used to fill the features in Assessment. Our assumption is that we do not lost important information for assessing the risk of complications.

3.2 Retrieval of similar cases

What we propose is to use CBR, particularly LID, to search for other patients having similar historical records. There are many authors proposing the [2, 15] use of CBR to manage diabetes. The system in [15] claims that the analysis and comparison of patterns of events can be more useful than just the analysis of single events as the other systems do. However, experiments does not support the idea that CBR is a good methodology since patients have different insulin metabolism rates and insulin tolerance levels, which influence the decision on the type and amount of insulin to be administered. We have to analyze this issue in detail, nevertheless our intuition is that the assessment of the complication risk, although it has some dependence from the patient metabolism, this is not so key as in the case of insulin dosage. In fact, we want to assess the general risk prognosis and this does not need to be so accurate as the determination of the insulin's type or the dose that a patient has to take.

Therefore, assuming that CBR is a good tool for our problem, what we have to determine now is how the retrieval of similar cases has to be done. First of all, patients with diabetes Type I have not to be retrieved as a precedents for a patient with diabetes Type II and vice versa since both types of diabetes are considered as different diseases by physicians. Also, the risk is also different between patients that already have some kind of complications and those that have not complications. These two considerations reduce the search for similar patients.

From here, the conceptual search is the one that takes into account all the series for all the features describing a patient. Although we do not want to take into account now the complexity of such search, we will have to face with other problems as the different length of the series. Therefore, we will have to study in depth techniques such as the ones used in SparseFGM [16]. Such system analyzes a series of lab test results of a potential diabetes patient to find particular complications that the patient may have. The goal is to diagnose diabetes complications from a set of lab test records of patients. In our approach, we do not only take into account lab tests but also all the features that the physician takes into account during a consultation. SparseFGM also takes into account the historical records of a patient but the results are based only on the records of that patient. In our approach we want to take into account the similar historical records to assess the complication risk of a patient.

We also think that we could use DIRAS as is now, reduce the search for similar patients. The idea is to find the features that are relevant to assess the risk of each complication. These relevant features in addition to other information we have, such as the years of diabetes development and the years of complications initiation could be key issues in searching the appropriate precedents.

3.3 Evaluation of Results

For each patient, we know the year of diabetes initiation and also the year of complications initiation. That is to say, we know the characteristics of the patient (analytical data, diet, lifestyle, etc) and also how many years have been elapsed from diabetes diagnosis to the initial complication. Therefore, we could use these dates to evaluate our approach: the prediction of initiation of development for each

complication should approximately coincide with the actual year of complications development.

ACKNOWLEDGEMENTS

This research is partially funded by the project RPREF (CSIC Intramural 201650E044) and the grant 2014-SGR-118 from the Generalitat de Catalunya.

REFERENCES

- [1] E. Armengol and E. Plaza, 'Lazy induction of descriptions for relational case-based learning', in *ECML-2001*, eds., L. De Raedt and P. Flach, number 2167 in Lecture Notes in Artificial Intelligence, pp. 13–24. Springer, (2001).
- [2] R. Bellazzi, C. Larizza, S. Montani, A. Riva, M. Stefanelli, G. d'Annunzio, R. Lorini, E.J. Gmez, E. Hernando, E. Brugus, J. Cermeneno, R. Corcoy, A. De Leiva, C. Cobelli, G. Nucci, S. Del Prato, A. Maran, E. Kilkki, and J. Tuominen, 'A telemedicine support for diabetes management: the t-idm project', *Int Clin Psychopharmacol*, **69**, 147–161, (2002).
- [3] K. Curran, E. Nichols, E. Xie, and R. Harper, 'An intensive insulinotherapy mobile phone application built on artificial intelligence techniques', *Diabetes Science and Technology*, **4**(1), 209–220, (2010).
- [4] M.J. Fowler, R. Bellazzi, C. Larizza, S. Montani, C. Pennati, C. Castelnovi, M. Stefanelli, M. Rondini, and R. Lorini, 'Telemedicine in the management of young patients with type 1 diabetes mellitus: a follow-up study', *Acta Biomedica*, **74**, Suppl. 1, 49–55, (2003).
- [5] M.J. Fowler, 'Microvascular and macrovascular complications of diabetes', *Clinical Diabetes*, **26**.
- [6] M. Hall and A.M. Felton, 'The st. vincent declaration 20 years on defeating diabetes in the 21st century', *Diabetes Voice*.
- [7] <http://www.diapedia.org/management/8105473810/the-st-vincent-declaration-on-the-treatment-of-diabetes>.
- [8] E. Keogh, J. Lin, and A. Fu, 'Hot sax: Efficiently finding the most unusual time series subsequence', in *Proceedings of the Fifth IEEE International Conference on Data Mining, ICDM '05*, pp. 226–233, Washington, DC, USA, (2005). IEEE Computer Society.
- [9] J.L. Kolodner, *Case-based Reasoning*, Artificial intelligence, Morgan Kaufmann Publishers, 1993.
- [10] C. Marling, M. Wiley, R. C. Bunesco, J. Shubrook, and F. Schwartz, 'Emerging applications for intelligent diabetes management', *AI Magazine*, **33**(2), 67–78, (2012).
- [11] A. Palaudàries, E. Armengol, and E. Plaza, 'Individual prognosis of diabetes long-term risks: A cbr approach', *Methods of Information in Medicine*, **40**, 46–51, (2001).
- [12] K. Piwernetz, 'Diabcare quality network in europe—a model for quality management in chronic diseases', *Int Clin Psychopharmacol*, **3**, 5–13, (2001).
- [13] M. Shokoohi-Yekta, Y. Chen, B. Campana, B. Hu, J. Zakaria, and E. Keogh, 'Discovery of meaningful rules in time series', in *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '15*, pp. 1085–1094, New York, NY, USA, (2015). ACM.
- [14] D. Trecroci, 'Insulin calculating, a tool that helps prevent errors', *Diabetes Health Magazine*, (2005).
- [15] A. Wills and I. D. Watson, 'Building a case-based reasoner for clinical decision support', in *PRICAI 2004: Trends in Artificial Intelligence, Auckland, New Zealand, August 9-13, 2004, Proceedings*, pp. 554–562, (2004).
- [16] Y. Yang, W. Luyten, L. Liu, M.F. Moens, J. Tang, and J. Li, 'Forecasting potential diabetes complications', in *Proceedings of AAAI'14*, pp. 313–319. AAAI Press, (2014).

D1NAMO, A Personal Health System for Glycemic Events Detection

Fabien Dubosson¹ and Stefano Bromuri² and Jean-Eudes Ranvier³ and Michael Schumacher⁴

Abstract. Several approaches are used nowadays to help diabetic people to handle their disease, one of them being the self-management of diabetes. We developed in this context a platform allowing patients to report and log their symptoms, medications and glucose levels through an Android application. In addition to self-management, the *DINAMO* project aims at using ECG signals in order to detect glycemic events and eventually predict glycemia levels. The BioHarness Zephyr 3 sensor has been integrated in the platform for this purpose. The resulting platform is a full-stack personal health system for diabetes self-management with support for physiological signals such as ECG: a physiological signals sensor, an Android application, a central server, a database and a few webpages are composing it. The question of the data lifecycle management in regards to the platform usages is discussed.

1 INTRODUCTION

The diabetes (diabetes mellitus) is a metabolic disorder characterized by chronic hyperglycemias — excessive glucose in the blood — due to defects in insulin level [1]. The type 1 diabetes includes causes due to a failure in the creation of the cells producing the insulin. The only treatment consists of taking insulin shots several times a day in order to regulate blood glucose level.

Several problems can arise from long-term diabetes, such as excessive risks of vascular diseases [2] or even damage, dysfunction and failure of various organs such as eyes and kidneys [1, 3]. Intensively controlled glycemia get type 1 patients to have a higher outcome on the risk of developing cardiovascular disease [3].

Insulin injections should be dosed correctly to avoid hypoglycemias — insufficient glucose in the blood — which are common side-effect of insulin therapy, especially for type 1 diabetes [6]. Severe hypoglycemias could be harmful for patients. This means there exists a trade-off between limiting the frequency of hypoglycemia while preventing cardiovascular disease later in patient's life.

The management of diabetes requires to take a drop of blood several times a day in order to measure the patient's glucose level. This measurement method is intrusive and the *DINAMO* project aims at exploring an alternative method using a non-intrusive measurement method that requires the collection of Electrocardiogram (ECG) data from patients in order to process them with machine learning algorithms. Such system would improve the quality of life of patients in two different ways. First by avoiding the patients to have to use intrusive measurement methods, and second by removing the need of

patients to think about checking regularly for hyper/hypo-glycemia, delegating this to an application that will throw alerts in such cases.

Up to our knowledge, no platform nor experiment to use the BioHarness' ECG in order to detect hypo/hyper-glycemia has been made yet. A review paper [4] explores the use of sensors to improve management of glucose and references two articles [10, 11] that are presenting methods that use the BioHarness, but only on Accelerometers and Heart Rate signals.

The presentation of the *DINAMO* project is made in the next section and the developed platform is described in the following one. A last section discuss the data lifecycle in regard to the platform usages.

2 DINAMO

The *DINAMO* acronym stands for *Diabetes type 1 Non-invasive Activity MONitoring* and aims at providing to type 1 diabetic patients a non-invasive way to manage their chronic disease. Several studies have shown that hypoglycemias are causing some modifications in the PQRST characteristics of ECGs, especially a prolongation in the QT intervals [5, 7, 8], as presented in Figure 1. One of these studies also suggests that this may allow the development of an hypoglycemia detection device [8]. The *DINAMO* project aims at using such technology to monitor type 1 diabetes in a non-invasive way.

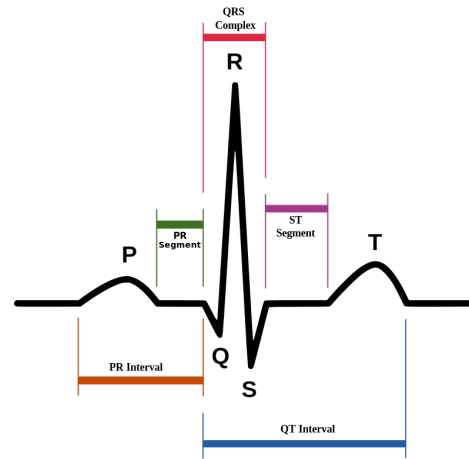


Figure 1. The PQRST characteristics with the QT interval

The *DINAMO* concept is the following: Diabetic patients are wearing an ECG sensor which is connected by Bluetooth to their smartphones. An Android application acts there as a controller to

¹ HES-SO Valais/Wallis, Switzerland, fabien.dubosson@hevs.ch

² Open University of the Netherlands, stefano.bromuri@ou.nl

³ EPFL, Switzerland, jean-eudes.ranvier@epfl.ch

⁴ HES-SO Valais/Wallis, Switzerland, michael.schumacher@hevs.ch

start/stop data transmission, as an helper to manage the disease by offering an interface to manually keep track of events, and as a buffer to store data while dealing with connectivity issues. The application send the data to a server that will analyze them on arrival, and then store them in a database for visualization. In case of a detected event, an alert is sent to the patient's phone, warning him about a potential event and asking him to take further measurements. Finally a web interface allows medical doctors to see their patients' data.

The studies having shown the prolongation of the QT interval have been made in a clinical setup by using medical-grade ECG devices. The *DINAMO* project does not fit in such category as it is based on a commercial sport-like chest belt for acquiring ECGs: The Zephyr BioHarness 3 shown in Figure 2. The feasibility of hypoglycemias detection in a real-life setup with a non-medical device is the goal of another part of the *DINAMO* project: some preliminary results with models description are presented in [9].



Figure 2. The Zephyr BioHarness sensor with its belt

The usual management of type 1 diabetes only requires patients to have a small pocket with them containing some needles, a stylus for needles, and a glucometer. The requirements for getting ECG data, as needed by the *DINAMO* project, are quite different: an ECG sensor and a smartphone. Additionally, the treatment of acquired ECG data requires a network connection on the phone in order to send the signals to a server, which will apply machine learning processing. Data are stored in a database by the server, and finally a web interface is needed to consult the data. The following section describes in more details all these components.

3 PLATFORM

The overall *DINAMO* platform is depicted in the Figure 3. This section describes in more details each component individually.

3.1 Sensor

The device that has been selected for *DINAMO* is the Zephyr BioHarness 3⁵. The selection has been made by a ponderation of different criterias such as price, ECG capabilities or connectivity. It is a sport-like chest-belt — shown in Figure 2 — that allows the acquisition of different kind of signals. It has three main sensors: ECG, Breathing, and Accelerometers; from which it is also able to extract higher level information. The data available over bluetooth are:

- ECG signal (250 Hz)
- Breathing signal (18 Hz)

⁵ <http://www.zephyranywhere.com/products/bioharness-3>

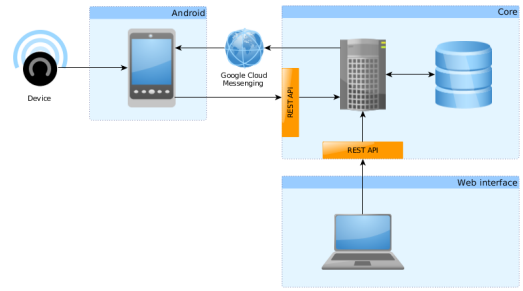


Figure 3. The overall platform architecture

- 3D Accelerometers signal (50 Hz)
- General information (1 Hz), among which:
 - Heart rate
 - Breathing rate
 - Posture
 - Activity level
 - Statistics like amplitude, noise, peaks, max or min about base signals

The device can be configured — over bluetooth — to send only the requested kinds of signals, meaning it is possible to optimize the battery life by requesting only the needed information.

3.2 Android

The sensor is connected by bluetooth to an Android application (Figure 4). The application asks the user to enable the bluetooth if not already done and offers a configuration menu to select the Bluetooth device to use. Another menu allows patients to select which packets should be sent from the device.

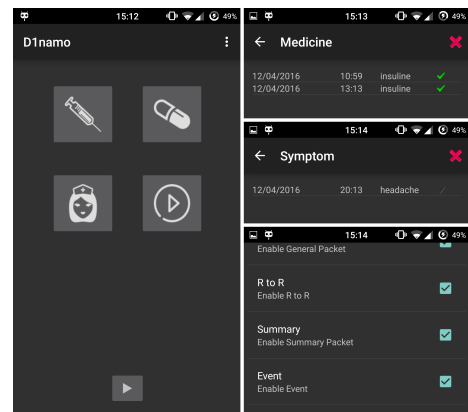


Figure 4. Some screens of the Android application

As the smartphone connectivity may be interrupted, the Android application has been designed to serve as a data buffer. This means that the data are not continuously sent over the network, but that the application gather the data locally before sending them as a batch on

a regular time interval, or when a given memory threshold has been reached. Another benefit of this approach is the battery saving that arise of not having the data channel open all the time.

The application also provides helping functionalities for diabetes management. Patients are offered interfaces for manually entering glucose measurement, medications and symptoms they may have taken/noticed. This can be seen as a personal diary allowing patients to discuss with the medical staff if the later notice anormal patterns in their signals.

3.3 Server

A central server gather the data from the Android application in order to process them by applying machine learning algorithms. The algorithms – worked out on another part of the *DINAMO* project [9] – will be integrated once performances would have been evaluated. The server is responsible to save the data inside a database in order to allow later visualization of the signals by the medical staff. For keeping the access to the data centralized, only the server is accessing the database but it provides an API to query the data.

The server application has been developed with Spring and JavaEE technologies on top of the Wildfly⁶ application server. Communication with the server are done through two different APIs, one allowing to receive data from the android application, and another one allowing to query data from the database. The communication through receiving API is not yet protected, but a placeholder library for encryption is already present in the pipeline. The decision on the encryption technology and algorithms still remains to be done.

3.4 Database

A PostgreSQL⁷ database is used to store the users physiological signals on the server side. A standard database table is used for storing users credentials, with a hashed and salted format for the password fields. The storage of data from the sensor is not done by saving one data per row as it is usually done, but instead by saving the data as gathered from the device in a bytes array format: the Zephyr sensor is using all bits of the packets sent over Bluetooth in order to minimize the energy needed. Saving the data in this format requires some processing for accessing data later on, so this may be changed in the future.

More generally, some discussions about the usefulness to keep all the records should be made with the medical staff. It should be possible to use some heuristics to discard records older than a given threshold age or to remove already seen data, with a feature to lock and prevent interesting ones to be removed.

3.5 Web interface

The platform currently comes with a few simple web pages allowing to manage the users (add, edit, delete) and to visualize users' data. The Figure 5 shows what the menu looks like.

All the features of the interfaces have not been implemented yet, but an evaluation of the usability of the existing web pages is planned. It will take the form of an qualitative evaluation with the medical staff and will lead the future development and enhancement of interfaces.

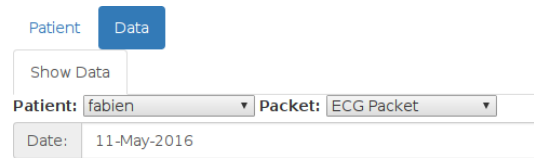


Figure 5. The menu of the web interface

3.6 Deployment

In order to allow an easy deployment of the different components, the docker⁸ software has been used. It allows to package binaries of applications with their files in a single entity called a “container”. Such container can be build in a reproducible and automated way, and it is possible to reuse existing containers of already packaged software. The PostgreSQL database for instance can be started from an official docker’s container, with a single command that will take care of fetching the container online and starting it. The server itself is provided as a docker container. Finally a “Makefile”⁹ orchestrates the lunch of the different containers to allow administrators to easily setup the whole platform.

4 DISCUSSION ON DATA LIFECYCLE MANAGEMENT

Signals such as ECG or Accelerometers output are acquired at high-frequency rates. The BioHarness 3 is getting the ECG signal at 250 Hz, while the Accelerometers are sampled at $3 \times 50\text{Hz}$ and the Breathing at 18 Hz. Storing such kinds of data in relational database tables will grow the number of entries quickly: summing these signals together, they represent 418 values per second, which adds up to more than 1,5 millions entries per hour. The data Acquisition that has been made for the project showed an usage of the device for at least 12h per day. Hence, an instance gathering the data of 20 patients, 12 hours a day during 1 month will accumulate more than 10 billions entries. It is possible to estimate the lower bound of space needed by the generated data. By using the device data sheet, we can get the precision of each kind of signals values, i.e. the number of bits that are used for each:

- ECG: $250 \text{ Hz} \times 10 \text{ bits} = 2500 \text{ bits/second}$
- Breathing: $18 \text{ Hz} \times 10 \text{ bits} = 180 \text{ bits/second}$
- Accelerometers: $50 \text{ Hz} \times 3 \text{ signals} \times 10 \text{ bits} = 1500 \text{ bits/second}$

Which leads to a total of 4180 bits per second, which is around 523 bytes. Using the same scenario as previously described, this sums up to more than 13 GB per month for 20 patients. While good relational databases can handle such high number of queries, and hard drives being cheap enough to handle the storage easily, this is not without raising up some questions about the data lifecycle management.

The different usages of the platform are triggering different needs in term of lifecycle management. Three classes of usages can be derived from the platform: the alerting need for patients, the querying and visualizing needs for the medical staff, and finally the machine learning need for researchers.

⁶ <http://wildfly.org/>

⁷ <http://www.postgresql.org/>

⁸ <http://www.docker.com/>

⁹ <https://www.gnu.org/software/make/>

The *patients* need is allowing patients to receive alerts regarding their blood glycemic state. This goal requires the last minutes of received data to be analyzed in order to detect glycemic events. The smartphone is sending data in a batch on regular intervals, so the analyze may be triggered on data reception before putting them into the database.

The *medical staff* needs are the visualization of patients data and the querying of past signals events. Both of these goals are in the target of relational databases as they are made for querying data, either it is for a visualization purpose or for finding event. This brings a first question about the data lifecycle: which data should be kept, for how long, and for which goal. However these questions are easy to address by discussing with the medical staff who can decide which kind of data they want to have, and for how long.

The *researchers* needs are to keep the data available for further research, and using incoming data for training algorithms. Creating backups of all signals for later use in research can be made easily by dumping the database. On the other side, depending on the machine learning techniques used, models refinements are possible. These could be done when data are arriving.

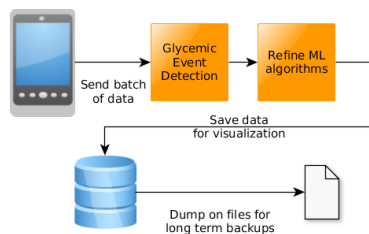


Figure 6. The data lifecycle

The data lifecycle management of such Personal Health Systems could then follow this schema (depicted in Figure 6): The data are sent as batch to the server. Data arrival trigger an analysis of the data in order to detect eventual glycemic events for the patient. The data can then be used to refine machine learning algorithms before being saved in the database. Database dumps could be done when data are needed, or when limits are reached. On a regular basis — that should be discussed with the medical staff — a cleanup of old data can be made to save space and avoid performance issue later on.

5 CONCLUSION AND FUTURE WORK

In this paper we present the platform we developed in the context of the *DINAMO* project. The platform allows diabetic patients to gather their physiological signals, such as ECG, Breathing or Accelerometers output, into a central database. Predictions about their glycemic states and detection of eventual glycemic events, such as hypo- or hyper-glycemias, can then be made out by using machine learning algorithms. The data lifecycle is also discussed in regards to the different usages of the platform.

By using the platform, medical doctors will be able to access and visualize their patients data. The developed user interfaces are in their first version, but a qualitative evaluation by a medical staff is planned in order to improve their usability. The detection of glycemic events is part of another side of the *DINAMO* project with some preliminary results, but formal performances evaluation still remains to be done.

Once the *DINAMO* project will be fully integrated, this platform will serve as a proof of concept for the validation of the feasibility of such non-invasive technologies in real conditions. This platform is not ready for production as several improvements should be made before being used outside of the research area, especially as medical platforms require a special care on security for users and data protection.

The future work on this platform includes the integration of the machine learning algorithms developed on the second part of the *DINAMO* project, as well as the integration of a query interface to allow the medical staff to search for patterns in the patients data.

ACKNOWLEDGEMENTS

This research has been financed by the *Nano-Tera.ch* initiative through the *DINAMO* project.

REFERENCES

- [1] Kurt George Matthew Mayer Alberti and PZ ft Zimmet, 'Definition, diagnosis and classification of diabetes mellitus and its complications. part 1: diagnosis and classification of diabetes mellitus, provisional report of a who consultation', *Diabetic medicine*, **15**(7), 539–553, (1998).
- [2] Emerging Risk Factors Collaboration et al., 'Diabetes mellitus, fasting blood glucose concentration, and risk of vascular disease: a collaborative meta-analysis of 102 prospective studies', *The Lancet*, **375**(9733), 2215–2222, (2010).
- [3] Diabetes Control, Complications Trial, et al., 'Intensive diabetes treatment and cardiovascular disease in patients with type 1 diabetes', *The New England journal of medicine*, **353**(25), 2643, (2005).
- [4] Sandrine Ding and Michael Schumacher, 'Sensor monitoring of physical activity to improve glucose management in diabetic patients: A review', *Sensors*, **16**(4), 589, (2016).
- [5] Bodil Eckert and Carl-David Agardh, 'Hypoglycaemia leads to an increased qt interval in normal men', *Clinical Physiology*, **18**(6), 570–575, (1998).
- [6] Graham P Leese, Jixian Wang, Janice Broomhall, Paul Kelly, Andrew Marsden, William Morrison, Brian M Frier, and Andrew D Morris, 'Frequency of severe hypoglycemia requiring emergency treatment in type 1 and type 2 diabetes a population-based study of health service resource use', *Diabetes care*, **26**(4), 1176–1180, (2003).
- [7] JLB Marques, E George, SR Peacey, ND Harris, IA Macdonald, T Cochrane, and SR Heller, 'Altered ventricular repolarization during hypoglycaemia in patients with diabetes', *Diabetic Medicine*, **14**(8), 648–654, (1997).
- [8] J Meinhold, T Heise, K Rave, and L Heinemann, 'Electrocardiographic changes during insulin-induced hypoglycemia in healthy subjects', *Hormone and metabolic research= Hormon-und Stoffwechselforschung= Hormones et metabolisme*, **30**(11), 694–697, (1998).
- [9] Jean-Eudes Ranvier, Fabien Dubosson, Jean-Paul Calbimonte, and Karl Aberer, 'Detection of hypoglycemic events through wearable sensors', *International Workshop on Semantic Web Technologies for Mobile and Pervasive Environments*, (2016).
- [10] Matthew Stenerson, Fraser Cameron, Shelby R Payne, Sydney L Payne, Trang T Ly, Darrell M Wilson, and Bruce A Buckingham, 'The impact of accelerometer use in exercise-associated hypoglycemia prevention in type 1 diabetes', *Journal of diabetes science and technology*, 1932296814551045, (2014).
- [11] Matthew Stenerson, Fraser Cameron, Darrell M Wilson, Breanne Harris, Shelby Payne, B Wayne Bequette, and Bruce A Buckingham, 'The impact of accelerometer and heart rate data on hypoglycemia mitigation in type 1 diabetes', *Journal of diabetes science and technology*, **8**(1), 64–69, (2014).

Ontologies for social, cognitive and affective agent-based support of child's diabetes self-management

Mark A. Neerincx^{1,2}, Frank Kaptein², Michael A. van Bekkum¹, Hans-Ulrich Krieger³, Bernd Kiefer³, Rifca Peters², Joost Broekens², Yiannis Demiris⁴ and Maya Sapelli¹

Abstract. The PAL project is developing: (1) an embodied conversational agent (robot and its avatar); (2) applications for child-agent activities that help children from 8 to 14 years old to acquire the required knowledge, skills and attitude for adequate diabetes self-management; and (3) dashboards for caregivers to enhance their supportive role for this self-management learning process. A common ontology is constructed to support normative behavior in a flexible way, to establish mutual understanding in the human-agent system, to integrate and utilize knowledge from the application and scientific domains, and to produce sensible human-agent dialogues. This paper presents the general vision, approach, and state of the art.

1 Ontologies in Cognitive Engineering

In Europe, an increasing number of about 140,000 children (<14 year old) have Type 1 Diabetes Mellitus (T1DM) [1]. The PAL project develops an Embodied Conversational Agent (ECA: robot and its avatar) and several applications for child-agent activities (e.g., playing a quiz and maintaining a timeline with the agent) that help these children to enhance their self-management (PAL, Personal Assistant for healthy Lifestyle, is an European Horizon-2020 project; www.pal4u.eu). In addition, it develops dashboards for caregivers (like diabetes nurses and parents) to enhance their supportive role. The general objective is to establish a smooth transition of the diabetes care responsibility from caregiver to the developing child, so that the child will have the required knowledge, skills, and attitude for adequate self-management at adolescence.

PAL is part of a joint, cognitive system, in which humans and agents share information and learn to improve self-management. The required sharing of (evolving) knowledge in the envisioned “blended care” setting has four important challenges:

1. To address the values & norms of both the caregivers in their different hospitals (e.g., diabetes regimes), and the caretakers in their different contexts (e.g., privacy, literacy).
2. To establish mutual understanding (a) within and between the different stakeholders of the PAL system (e.g., the end-users like children and caregivers and research & developers like academics and engineers), and (b) between the humans and PAL-agents.

3. To continually acquire, utilize and deploy knowledge about child's self-management support.
4. To produce natural, flexible, personalized human-agent interactions that are sustainable in the long term as well as allow to extract data about the user from these interactions.

To meet these four challenges, we are developing an ontology as an integrated part of system development, i.e., in a systematic, iterative, and incremental cognitive engineering process. First, available ontologies and approaches are assessed and, possibly, improved and integrated for our purposes (section 2). Second, relevant theories and models of the concerning scientific research fields are identified and formalized for adoption in the ontology (section 3). Third, the ontology is implemented in an artefact or prototype (i.e., the PAL system) and, subsequently, tested and refined (section 4).

2 Models for Diabetes Self-Management

Because PAL covers a large domain of interest, we have developed ontology models as high-level building blocks for smaller, separate areas of interest (frames). First, appropriate frames were selected from existing (global) libraries and, if needed, tailored to the PAL purposes. Second, for the missing elements, frames were modeled by constructing a new ontology. Subsequently, the individual frame models were related (interlinked) in an integrated PAL model. Because most existing ontologies provide “only” a partial fit to the intended scope of PAL, we needed to adapt these models by extending them (e.g., when concepts were lacking), or by selectively downsizing them (e.g., when there were too many details or concepts in the model). The frames we have identified and modeled so far are among others: (1) human and machine roles involved in self-management; (2) emotions and sentiments that cover the emotional responses of both robot and child to interaction as well as the general state of mind of the child; (3) tasks that include among other things: learning and self-management tasks, associated goals, and objects; (4) issues related to medical examinations (e.g., lab values); and (5) dialogue management through a combination of dialogue acts and shallow semantics. A more elaborate PAL ontology will also include interaction and behavior models of robot and avatar, a model for privacy of information of self-management activities and a model to cover the agreements and social contracts between child and ECA.

Figure 1 provides a simple example of the task frame (cf. [2]). An Agent, such as a child or avatar, is an entity that performs a certain task, like an educative quiz game. An associated goal

¹TNO, The Netherlands, mark.neerincx@tno.nl

²Delft University of Technology, Netherlands, F.C.A.Kaptein@tudelft.nl

³DFKI, Germany, krieger@dfki.de

⁴Imperial College, United Kingdom, y.demiris@imperial.ac.uk

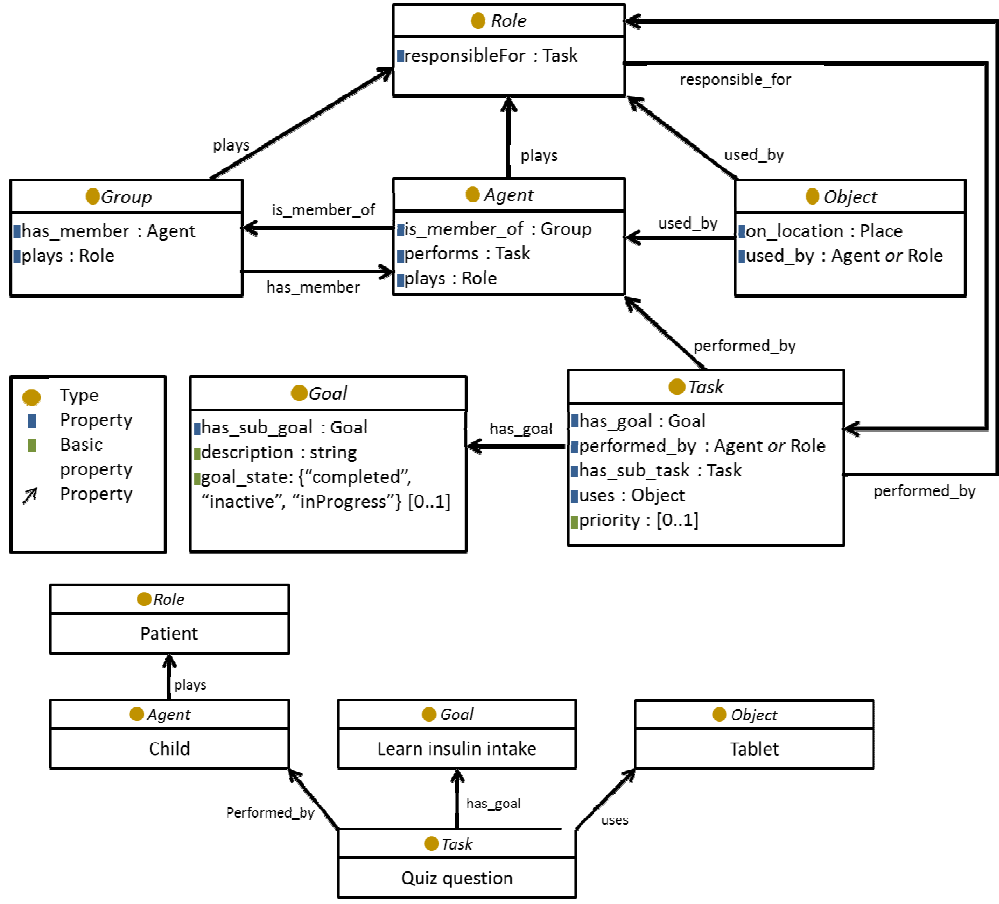


Figure 1: Simple example of the general task frame at the top and an instantiation at the bottom.

(e.g., learning about Insulin taking) can be attained by performing the related task (e.g., answering related questions correctly while playing the quiz). Objects such as a tablet device, are typically used when performing the task. The agent has a role while performing the task (e.g., patient) and can be part of a group of agents (e.g., parents).

Important objectives of the PAL ontology are norm-compliance, shared understanding, interpretation, reasoning, and generation of verbal utterances. The ontology is based on a uniform representation of an application semantics that uses dialogue acts and frames that are defined in an extended RDF and OWL ontology [3]. In addition, all data that influence multimodal utterance generation are specified in the ontology (e.g., user data), which facilitates access and combination of the different bits of information. We heavily extended existing processing components, e.g., the reasoning engine HFC from DFKI and its database layer [4], which make information available to the interaction management and analysis. We defined a new formalism for the specification of dialogue policies that combines dialogue rules, transaction time-based knowledge representation [5], and dialogue history in a unique way. One important part of the PAL ontology combines dialogue acts using the DIT++ standard [6] and semantic frames, loosely based on thematic relations [7], used in today's frameworks VerbNet, VerbOcean, or FrameNet. Below, we show a

simplified version of the combined representation, built for the sentence: "I could show you a picture of the last football game".

Offer(Showing, theme=Picture, sender=I MYSELF, addressee=NAO ROBOT, topic=Football).

3 Integrating Relevant Theories

In the PAL project, dedicated studies of models in the concerning scientific research areas are being conducted. For supporting the *social processes* that are involved in self-management learning, PAL models relationships in terms of familiarity or intimacy, liking, attitude and benevolence [8]. Particularly, the child-ECA bonding process is being supported by the Dyadic Disclosure Dialog Module (3DM) that supports the mutual child-agent self-disclosures. The PAL ontology distinguishes three main classes for these dialogues: disclosure, prompt and closer. In addition to valence and topic, each disclosure has an intimacy level according to the 4-level Disclosure Intimacy Rating Scale (DIRS). Burger et al. (2016) provide more detailed information on the 3DM of PAL and its theoretical foundations [9].

For supporting the *cognitive processes*, the diabetes knowledge and corresponding learning goals have been modeled to monitor and reason about progress (e.g., on diabetes regimes, self-control,

food, physical exercises, and stress coping). Goal attainment is an important indicator of the changes in behavior of children [10], and can be supported by personalized feedback of the ECA. Figure 2 provides a simplified sketch of a dialogue instantiation in the PAL system. Answering a quiz question is an example of a task (Fig 1). Answering correctly (partly) fulfills one or more (learning) goals. Note that the same goal can be satisfied by another task too, such as a sorting game. The different goals have specific difficulty levels (0-3). The caregivers decide what goals are currently relevant and achievable for a child. Together with caregivers, a child selects the specific goals to attain: `<child:URI> <hasGoal> <goal:URI>`. Since the system will only suggest tasks that can achieve the child's current goals, these tasks are implicitly following these same difficulty levels. For example, a quiz question that satisfies a level 3 goal will be more difficult than a question satisfying a level 0 goal. Goal attainment is an important aspect of self-management. PAL will monitor the goal attainment progress: `<Goal:URI> <hasProgress> float`. For every goal, the ontology defines what tasks, and (sub-)goals should be achieved to achieve the goal itself. GoalProgress is function of goal:neededForAsClass and goal:requiresAsClass. By computing the percentage of tasks, subtasks, and sub-goals currently achieved, the system computes a current progress on this goal. This is recorded with a time stamp, so that progress over time can be calculated.

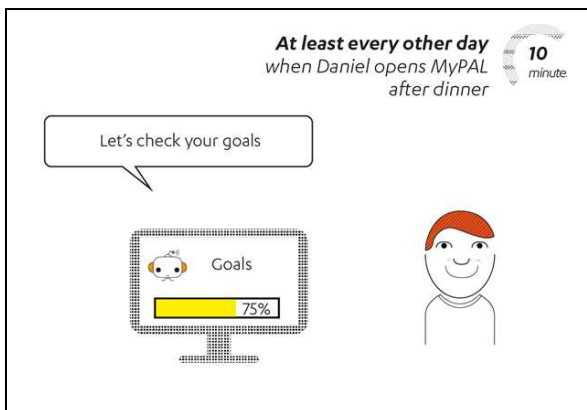


Figure 2: Simplified situated speech act of the avatar.

For supporting the *affective processes*, the PAL system introduces several methods to model the affective state of a child. First, sentiment mining technology is applied to estimate child's mood in the child-PAL textual dialogues [11]. Second, in the tablet application, the child can further self-report on the experienced emotions and moods for activities the child performed during the day. Third, the child model will estimate emotions experienced by the child resulting from activities proposed by the ECA. For example, the ECA can propose to play a quiz with the child, and predict joy when the child did well during the quiz. This child model is based on the belief desire theory of emotions [12, 13], in which emotions are a direct consequence of beliefs and desires of an individual. For example, if one believes X and desires X, then one is happy about X. This way, the child model can reason about the child's beliefs and desires. The model improves over time. If the child self-reports positive emotions during an activity while the child model estimates negative ones, then the child model updates the beliefs-desire assumptions concerning the child. The PAL ontology will represent complex affective states. Emotions are

directed at objects, or events, and are short intense episodes. Moods are undirected and less intense, but linger for a prolonged period of time. Emotions are stored with the activity that had this emotion as a consequence. Moods contain a timestamp, indicating when it was measured. This representation makes it possible to find correlations between activities and affect over a prolonged period of time.

4 Implementation and Evaluation

The PAL system consists of several modules with dedicated support objectives. For example, the dialogue manager aims at engaging conversations between child and the ECA, the action-selection module HAMMER [14] learns over time what the best actions are (e.g., proposing to play a quiz, or starting a new dialogue) to improve the child's knowledge of diabetes while maintaining a positive emotional state for the child, and the child model aims at estimating the emotional states.

Figure 3 shows the data flows of the PAL system with an extendable set of modules that communicate through a common Nexus. When a module has new information to share with other modules (e.g., action selection proposes to play a quiz) then this information is posted on the Nexus. Any module can read and use this new information. The application can then read this proposal and start a quiz on the tablet, and/or the dialogue manager can start a small dialogue by asking the child whether he/she wants to play a quiz. The PAL ontology provides the shared knowledge representations, defined in the extended HFC reasoner and allowing for testing and refining.

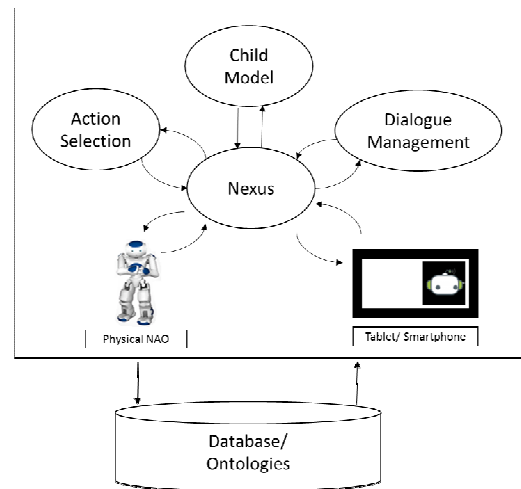


Figure 3: The PAL system.

Currently, we are analyzing the first data sets of children and caregivers that used the PAL system in diabetes camps, hospitals and at home (in Italy and in the Netherlands) from a few days to 4 weeks. Based on the ontological concepts, we can identify meaningful patterns in the data that will be used to improve the intelligence of PAL, e.g. concerning the goal attainment progress (i.e., enhance the knowledge base with refined ontology and reasoning mechanisms). Furthermore, the data analysis will help to refine the ontology substantially. For example, parents' relationship (cohabit or divorce) seems to affect child's PAL usage (quantity and regularity) substantially. These concepts with their

mutual relations are being added to the ontology to “feed” mitigating support functions. A second example concerns the identified cultural differences in Italian and Dutch children for the wealth and directness of their multimodal interactions with the robot [11]. Among other things based on these results, the child and robot models will be enriched to establish adaptive — personalized and culture-harmonized— child-robot interactions.

5 Discussion

The PAL project develops personalized support for children, helping them to acquire the required attitude, knowledge and skills for adequate diabetes self-management. It applies a situated Cognitive Engineering (sCE) methodology to design and test: (1) an ECA for children, (2) several (educative) child-ECA activities, and (3) dashboards for caregivers. This methodology includes an ontology engineering component to establish a system’s knowledge base that is univocal, theoretically sound, coherent, consistent and transparent [15]. The resulting common ontology is used to establish mutual understanding in the human-agent system, to integrate and utilize knowledge from the application and scientific domains, and to produce sensible human-agent dialogues. For the first version of the PAL ontology, a network of connected ontologies (“frames”) have been constructed, each consisting of general concepts and their relations. The “dialogue management frame” was worked out in more detail, i.e., the specification of the data structures to be used by the dialogue specifications, dialogue history, and information state. Furthermore, the reasoning components were adapted, so that this knowledge source can be used efficiently once the formalism specification is fully implemented.

The PAL project entails multi-disciplinary research and design of a “blended care” system with the involvement of a large diversity of stakeholders. In general, the ontology construction helped to identify (interrelated) key concepts that should be univocally addressed in the design (e.g., requirements), implementation (e.g., dialogues) and evaluations (e.g., goal attainment). Furthermore, it enforces the systematic integration of relevant theories on social, cognitive and affective processes into the support system (e.g., on bonding, goal-driven learning and emotion). In line with the general iterative development process, the ontology will be refined for enhanced self-management support in the next versions of the PAL system.

It is interesting to note that the PAL ontology can be viewed as a frame-based ontology in terms of Minsky [16] and Hoekstra [17]: An explicit, structured, and semantically rich representation of declarative knowledge like psychological theories of human cognition use, distinguishing “frames” or “classes” (upper level) from “instantiations” (lower level). This approach seems therefore particularly appropriate for representing knowledge involved in learning [15], e.g., learning to cope with a chronic disease.

Acknowledgements

PAL is funded by Horizon2020 grant nr. 643783-RIA.

References

- [1] Freeborn, D., Dyches, T., Roper, S.O. and Mandelco B. (2013). Identifying challenges of living with type 1 diabetes: child and youth perspectives, *Journal of clinical nursing*, vol. 22, no. 13-14, pp. 1890–1898.
- [2] Van Welie, M., Van der Veer, G.C. and Eliëns, A. (1998). An ontology for task world models. *Eurographics Workshop on Design Specification and Verification of Interactive Systems*, pp.3–5
- [3] ter Horst, H. J. (2005). Completeness, decidability and complexity of entailment for RDF Schema and a semantic extension involving the OWL vocabulary. *Journal of Web Semantics*, 3:79–115.
- [4] Krieger, H.-U. (2013). An Efficient Implementation of Equivalence Relations in OWL via Rule and Query Rewriting. *Proceedings of the 7th International Conference on Semantic Computing (ICSC)*.
- [5] Krieger, H.-U. (2016). Capturing Graded Knowledge and Uncertainty in a Modalized Fragment of OWL. *Proceedings of the 8th International Conference on Agents and Artificial Intelligence (ICAART)*.
- [6] Bunt, H., Alexandersson, J., Choe, J.-W., Fang, A.C., Hasida, K., Petukhova, V., Popescu-Belis, A., and Traum, D. (2012). ISO 24617-2: A semantically-based standard for dialogue annotation. *Proceedings of the 8th International Conference on Language Resources and Evaluation (LREC)*.
- [7] Fillmore, C.J. (1977). The Case for Case Reopened. In: *Grammatical Relations. Syntax & Semantics*. Academic Press.
- [8] Altman, I. and Taylor, D. (1973): *Social penetration theory*. Holt, Rinehart & Mnston, New York.
- [9] Burger, F., Broekens, J. and Neerincx, M.A. (2016). A Self-Disclosing Companion Agent for Children. *Proc. of the Intelligent Virtual (IVA) 2016 Conference*. Springer Int. Publishing
- [10] Kleinrahm, R. Keller, F., Lutz, K. Kölch, M. and Fegert, J.M. (2013). Assessing change in the behavior of children and adolescents in youth welfare institutions using goal attainment scaling. *Child & Adolescent Psychiatry & Mental Health*, 7:33 (11 pages).
- [11] Neerincx, A., Sacchitelli, F., Kaptein, R., van der Pal, S., Oleari, E., & Neerincx, M. A. (2016). Child’s Culture-related Experiences with a Social Robot at Diabetes Camps. Eleventh ACM/IEEE International Conference on Human Robot Interaction (pp. 485-486). IEEE Press.
- [12] Reisenzein, R. (2001). Appraisal processes conceptualized from a schema-theoretic perspective: Contributions to a process analysis of emotions.
- [13] Reisenzein, R. (2009). Emotions as metarepresentational states of mind: Naturalizing the belief–desire theory of emotion. *Cognitive Systems Research*, 10(1), 6-20.
- [14] Demiris, Y. and Khadhour, B. (2006). Hierarchical Attentive Multiple Models for Execution and Recognition (HAMMER), in *Robotics and Autonomous Systems*, 54:361-369.
- [15] Peeters, M. M., Bosch, K. V. D., Neerincx, M. A., & Meyer, J. J. C. (2014). An ontology for automated scenario–based training. *International Journal of Technology Enhanced Learning*, 6(3), 195-211.
- [16] Minsky, M. (1975). *A framework for representing knowledge*. The Psychology of Computer Vision, McGraw-Hill, New York.
- [17] Hoekstra, R. (2009). *Ontology Representation: Design Patterns and Ontologies that Make Sense*, IOS Press.

Handling Missing Phenotype Data with Random Forests for Diabetes Risk Prognosis

Beatriz López,¹ Ramon Viñas,² and Ferran Torrent-Fontbona,³ and José Manuel Fernández-Real⁴

Abstract. Machine learning techniques are the cornerstone to handle the amounts of information available for building comprehensive models for decision support in medical practice. However, the datasets use to have a lot of missing information. In this work we analyse how the random forests technique could be used for dealing with missing phenotype values in order to prognosticate diabetes type 2.

1 INTRODUCTION

Diagnosis of type 2 diabetes is made typically using clinical criteria. However, some population studies, specially in which young people is involved, have provided evidence that the diagnosis should be supported by phenotype data [16]. This phenotype data is not just useful for handling inheritance factors, but also for understanding nutrition conditions in pre and post-natal stages (see [8] and [9] for a reviewed version). In fact, phenotype data could provide new possibilities for handling risk prognosis for both, type 1 and type 2 diabetes [17], and also find explanations for other combination processes known as undetermined diabetes or 1.5 diabetes [16].

Our work concerns on using phenotype data to building a clinical decision support system (CDSS) for diabetes 2 prognosis. To that end, we are provided with a huge dataset of patient samples, each one characterised by a considerable amount of phenotypes. Therefore, we require the application of a machine learning technique to obtain a prognosis model to be handled by the CDSS. In so doing, our challenge is to handle the considerable amount of missing information, a typical situation when dealing with phenotypes [14].

There are several methods to deal with missing data that can be organized in four categories [15]. First, methods that discard instances (i.e. samples) with missing information. Second, methods that acquire missing values to complete the information, which involves some additional costs. Third, imputation methods are the largest family, and can be in turn organized in three groups: predictive value computation methods (e.g. mean, mode, the most popular ones), distribution-based computation (which take into account the class or diagnose of the samples), and unique-value imputation (replacing the missing value by a given value that represents it). Finally, the fourth category of methods are the reduced-feature models which incorporate only the phenotypes known in a given query (test). These latter kind of methods have been shown to be the ones that most improve the prognosis accuracy [15]

Handling missing values by adding and removing features according to a given query as reduced model approaches do is quite similar to the random forests (RF) machine learning technique. RF is a method that combines several decision tree models to provide a classification outcome (i.e. prognosis) [5]. Each decision tree is learned by using a base learner method applied to a subset of features (phenotypes) that are randomly selected, as well as to a subset of samples that are also randomly chosen. In fact, the RF technique could be considered as a combination of discard instance methods and reduced-feature models for handling of missing values. However RF does not remove any information which could be useful towards a personalised prognosis. In this paper, we analyse such possibility by applying RF to prognosticate diabetes type 2 from a dataset of phenotypes with a considerable amount of missing values.

This paper is organized as follows. First, we describe in Section 2 some previous related work. Next, in Section 3 we explain our method. We continue in Section 4 by describing the experimentation carried out and providing the results obtained. We end the paper in Section 5 with some conclusions and discussion about future work.

2 RELATED WORK

The application of machine learning techniques to gene expression data is becoming a key issue for Biomedicine [3]. For example, [7] build a binary logistic regression model based on phenotypes and genotype data to risk prediction of inheritance diabetes. 5639 patients were considered in the study, from which samples with at most a 10% of missing features were considered. We are not provided with so many patient data, and we need to handle a higher number of missing information to keep enough samples for learning a model.

In [14] and approach for imputing missing phenotypes based on a method called co-trained is presented. Co-trained means that missing phenotypes are predicted (*in-silico* phenotypes) based on a second class of information (i.e. clinical data). The method is applied in phenotypes related to migraine. the use of in-silico phenotypes generation implies that two machine learning methods are combined (one for phenotype learning, the second one for disease prediction from the phenotypes), and transfer learning complex issues should be taken into account. Our aim is to keep original data as much as possible, handling missing data in the machine learning technique itself.

Another interesting work is [11], which use self-organizer maps to look for associated diseases (kidney disease, retinopathy, hypertension). Self organized maps allows to obtain groups of biomarkers than should next be interpreted by the clinicians. In our work, we are dealing with classification (i.e. prognosis), although [11] could be considered to extend the follow-up of diagnostic persons, in a hybrid methodology of [11] and ours.

¹ University of Girona, email: beatriz.lopez@udg.edu

² University of Girona, email: rvinast@gmail.com

³ University of Girona, email: ferran.torrent@udg.edu

⁴ Biomedical Research Institute of Girona, email: jmfreal@idibgi.org

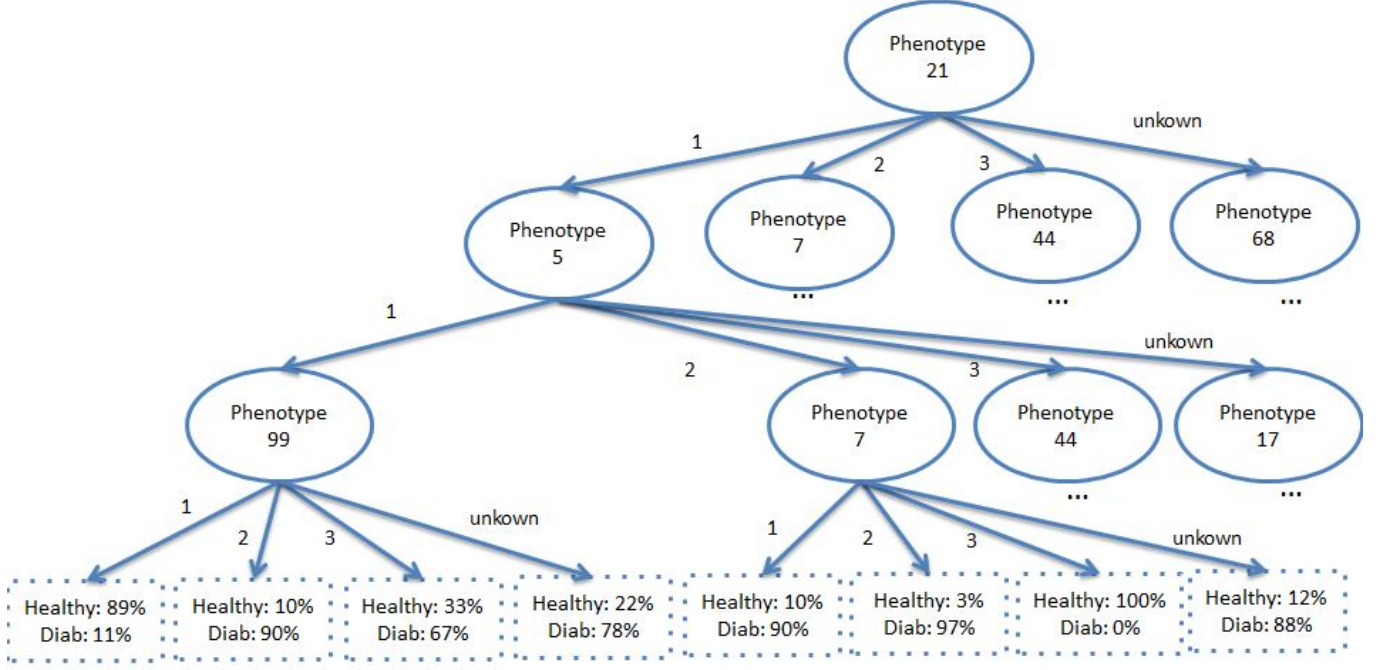


Figure 1. Example of Random Forests.

In [1] a comparison analysis among different imputation methods is performed, including instance deletion, mean imputation, median imputation, and k-nearest neighbour (knn) over a parametric and a non-parametric machine learning methods. The results highly depend on the characteristics of the data set, that is, the amount of missing features. Nevertheless, it seems that the case-deletion methods is the one that performs the worst, while the knn showed a higher robustness to missing data. The latter results agree with [2], where the authors analyse also several methods and demonstrate the out-performance of knn. The knn approach was analysed also in [15] as part of the reduced model approaches, and the results were slightly different, obtaining best performance with the authors approach called reduced-feature ensemble (RFE). RFE consists on generating several models, in which a feature is excluded in each of them. Given a query case, the outcomes of the different models are combined in a voting approach to obtain the final prediction value. This approach is also known as bagging ("bootstrap aggregating") [4]. However, bagging suffers from a higher correlation of the predictions [12]. The RF technique applied in our work decorrelate the base learners thanks to the random choice of features and samples.

3 METHODOLOGY

Our aim is to build a prediction model from phenotype data, which involves a considerable amount of missing values. The technique we are proposing is RF, because our hypothesis is that RF are able to handle missing information in a similar way than remove-feature and remove-instance missing information methods. However, RF does not discard any data a priori, which could provide nice properties regarding individualization (i.e. personalized prognosis).

RF is a supervised method, meaning that each instance or sample is labelled with the outcome (prognosis). Each instance is noted as

(x, y) , where x is a list of attributes a_1, a_2, \dots, a_n and its values v_1, v_2, \dots, v_n ; and y the class to which the patient belongs. In our particular case, $y \in C = \{healthy, diabetesType2\}$. Moreover, a_i are the phenotypes, and we use v_{ij} to denote the j value of the i phenotype. Each phenotype i has NVA_i values. In our particular case, $NVA_i = 4$ ($\forall i$), 3 values, plus the *unknown* value. Therefore, we are considering phenotypes with missing information in our machine learning technique⁵.

RF consists of an ensemble of k classifiers $h_1(x), h_2(x), \dots, h_k(x)$, being $h(x)$ the joint classifier [13, 5]. Each classifier $h_i(x)$ consists of a decision tree, in which nodes are attributes (see Figure 1). The selection of which attribute is collocated in a node is performed as follows: 1) by randomly selecting a subset of features, 2) an evaluation measure is applied to the selected attributes according to their capability to provide homogeneity partitions of the samples, and 3) the attribute with the highest score is chosen. In particular, we use the change of the Gini impurity function (GC) to compute the score, as described in Equation 1:

$$GC(a_i) = - \sum_{C_k \in C} p^2(C_k) + \sum_{j=1}^{NVA_i} p(v_{i,j}) \sum_{C_k \in C} p^2(C_k|v_{i,j}) \quad (1)$$

Once a node is set with an attribute a_i , the data is split into as many sets as values the a_i attribute has. Then, the tree is growth with new nodes in each branch that are obtained by repeating the attribute selection process. The stopping conditions is defined according to the number of instances remaining in a node: if this number is lower than a given threshold τ , the algorithm stops. Samples used to build each tree are also selected randomly with replacement.

⁵ In fact, this could be considered as a unique-value imputation method, as the unknown or missing value is treated as another attribute value.

Given a query case q , each decision trees provides an outcome, $h(q)$, and the final prediction is obtained by using a voting mechanism.

4 RESULTS AND DISCUSSION

In this section we describe our data, the experimental scenarios, and the results obtained.

4.1 Dataset description

The experimentation has been carried out with a dataset of 1074 patients, of whom we knew whether they had diabetes or they do not. For 196 patients, the diagnosis was unknown and therefore, have been removed from the dataset, remaining a total of 878 instances for experimentation. Each sample contains 101 phenotypes regarding diabetes type 2.

Regarding missing information, Figure 2 shows the distribution of missing data along the different samples. It is worthy to observe that some of the samples accumulates a huge percentage of missing information. On the other hand, Figure 3 shown the amount of missing values per phenotypes⁶ (blue color). Phenotypes have been ordered in the x-axis according to their amount of missing values.

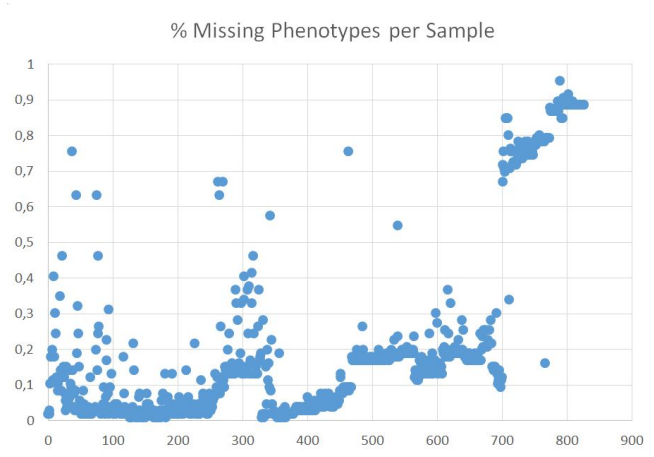


Figure 2. Percentage of missing phenotype values per sample. X-axis: cases.

4.2 Experimental set up

In order to analyse the implications of RF to handle missing data, the following experimental scenarios have been defined:

Raw data The dataset is used as provided.

Reduced features Features with the highest degree of missing information are removed. In particular, all features with more than 23% of missing values have been removed. This percentage has been set up according to the information visualized in Figure 3.

Reduced samples Samples with more than 25% of missing information has been removed. The percentage has been set up according to Figure 2.

⁶ Phenotypes names are hidden for simplicity reasons and medical research confidentiality issues.

Reduced features and samples Both, the reduced features and samples criteria is applied to the dataset.

The number of decision trees has been set to $k=1000$. According to [5], as the number of trees increases, for almost surely the RF converges to the real predictor. The experimentation methodology used has been the stratified k-fold cross validation (we set 5 folds). Results are analysed in terms of accuracy.

4.3 Results

Table 1 shows the results obtained in the different scenarios. The highest accuracy is obtained when removing samples with a huge amount of missing values (in bold). On the other hand, it is interesting to observe that the results when removing features are very bad, even when the removed features contain a lot of missing values. This fact also impacts in the combination scenario. Therefore, RF is handling appropriately missing information. Internally, RF are building several trees in which the phenotypes with a high amount of missing features could be skipped, but the presence of all of the phenotypes are important for prognosis prediction. In that regard, individualization is keep in the model, favouring a personalized prognosis.

On the other hand, RF is not able to handle samples with a huge number of missing information (scenario raw data). Although internally samples are randomly selected for building the decision trees, RF require from some pre-processing that filter outs the data with a huge amount of missing information in order to provide good accuracy results. Therefore, a pre-processing step for performing such remove-instances method is still required.

Table 1. Accuracy results

Scenario	Experiment	Accuracy
1	Raw data	80.50%
2	Reduced features	62.93%
3	Reduced samples	86.91%
4	Combine 2+3	62.67%

5 CONCLUSION

The application of machine learning techniques to phenotype datasets for building models for disease prognosis need to deal with a huge amount of missing information. In this work we present an application of RF that shows how this technique could deal with missing information. Results show than RF can perform well with features with missing values. Keeping all phenotypes lead us to think that RF favours personalized prognosis, considering all the particularities of an individual. However, regarding samples, RF requires a minimum information in the samples to achieve good accuracy results.

As a future work, we need also to explore the combination of phenotype data with clinical information, as well as other environmental factors; diabetes type 2 is an heterogeneous disorder that require considering all these factors [10]. On the other hand, the use of RF causes a loss of the nice interpretation properties of a single decision tree. In that regard, the work of [6] could provide some insights.

Acknowledgment

This project has received funding from the grant of the University of Girona 2016-2018 (MPCUdG2016) and the European Unions Hori-

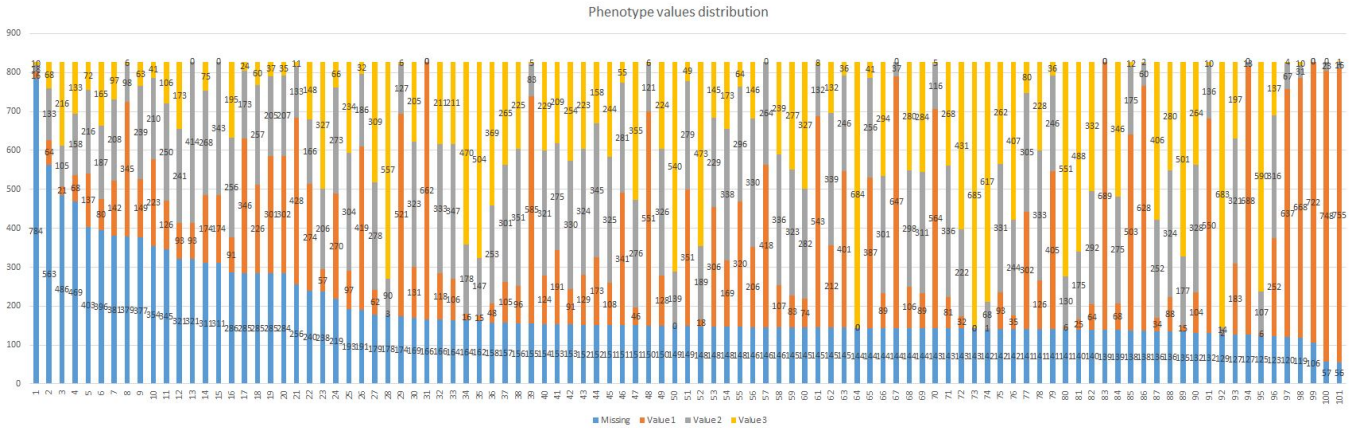


Figure 3. Distribution of phenotype values. Phenotypes are ordered according to the highest to lowest number of missing values (blue color).

zon 2020 research and innovation programme under grant agreement No 689810 (PEPPER). The work has been developed with the support of the research group SITES awarded with distinction by the Generalitat de Catalunya (SGR 2014-2016).

REFERENCES

- [1] Edgar Acuña and Caroline Rodriguez, ‘The Treatment of Missing Values and its Effect on Classifier Accuracy’, in *Classification, Clustering, and Data Mining Applications*, 639–647, Springer Berlin Heidelberg, Berlin, Heidelberg, (2004).
- [2] Gustavo E A P A Batista and Maria Carolina Monard, ‘An analysis of four missing data treatment methods for supervised learning’, *Applied Artificial Intelligence*, **17**(5-6), 519–533, (2003).
- [3] Riccardo Bellazzi and Blaz Zupan, ‘Towards knowledge-based gene expression data mining.’, *Journal of biomedical informatics*, **40**(6), 787–802, (dec 2007).
- [4] Leo Breiman, ‘Bagging Predictors’, *Machine Learning*, **24**(2), 123–140, (1996).
- [5] Leo Breiman, ‘Random Forests’, *Machine Learning*, **45**(1), 5–32, (2001).
- [6] Hugh A. Chipman, Edward I. George, and Robert E. McCulloch, ‘BART: Bayesian additive regression trees’, *The Annals of Applied Statistics*, **4**(1), 266–298, (mar 2010).
- [7] Hüsametdin Gül, Yeim Aydin Son, and Cengizhan Açikel, ‘Discovering missing heritability and early risk prediction for type 2 diabetes: a new perspective for genome-wide association study analysis with the Nurses’ Health Study and the Health Professionals’ Follow-Up Study.’, *Turkish journal of medical sciences*, **44**(6), 946–54, (2014).
- [8] C. N. Hales and D. J. P. Barker, ‘Type 2 (non-insulin-dependent) diabetes mellitus: the thrifty phenotype hypothesis’, *Diabetologia*, **35**(7), 595–601, (jul 1992).
- [9] C. N. Hales and D. J P Barker, ‘Type 2 (non-insulin-dependent) diabetes mellitus: the thrifty phenotype hypothesis’, *International Journal of Epidemiology*, **42**(5), 1215–1222, (2013).
- [10] H E Lebovitz, ‘Type 2 diabetes: an overview.’, *Clinical chemistry*, **45**(8 Pt 2), 1339–45, (aug 1999).
- [11] Ville-Petteri Mäkinen, Carol Forsblom, Lena M Thorn, Johan Wadén, Daniel Gordin, Outi Heikkilä, Kustaa Hietala, Laura Kyllönen, Janne Kytö, Milla Rosengård-Bärlund, Markku Saraheimo, Nina Tolonen, Maija Parkkonen, Kimmo Kaski, Mika Ala-Korpela, Per-Henrik Groop, and FinnDiane Study Group, ‘Metabolic phenotypes, vascular complications, and premature deaths in a population of 4,197 patients with type 1 diabetes.’, *Diabetes*, **57**(9), 2480–7, (sep 2008).
- [12] Kevin P. Murphy, *Machine Learning: A Probabilistic Perspective*, The MIT Press, 2012.
- [13] Marko Robnik-Sikonja, ‘Improving random forests’, *Machine Learning: ECML 2004*, **12**, (2004).
- [14] Damian Roqueiro, Menno J Witteveen, Verner Anttila, Gisela M Terwindt, Arn M J M van den Maagdenberg, and Karsten Borgwardt, ‘In silico phenotyping via co-training for improved phenotype prediction from genotype.’, *Bioinformatics (Oxford, England)*, **31**(12), i303–10, (jun 2015).
- [15] Maytal Saar-Tsechansky and Foster Provost, ‘Handling Missing Values when Applying Classification Models’, *The Journal of Machine Learning Research*, **8**, 1623–1657, (2007).
- [16] Hala Tfayli, Fida Bacha, Neslihan Gungor, and Silva Arslanian, ‘Phenotypic type 2 diabetes in obese youth: insulin sensitivity and secretion in islet cell antibody-negative versus -positive patients.’, *Diabetes*, **58**(3), 738–44, (mar 2009).
- [17] Tiinamaija Tuomi, ‘Type 1 and type 2 diabetes: what do they have in common?’, *Diabetes*, **54 Suppl 2**(suppl 2), S40–5, (dec 2005).